The National Children's Study 2014

AN ASSESSMENT

Panel on the Design of the National Children's Study and Implications for the Generalizability of Results

Greg J. Duncan, Nancy J. Kirkendall, and Constance F. Citro, *Editors*

Committee on National Statistics

Division of Behavioral and Social Sciences and Education

Board on Children, Youth, and Families

Institute of Medicine

NATIONAL RESEARCH COUNCIL *AND*
INSTITUTE OF MEDICINE

THE NATIONAL ACADEMIES PRESS
Washington, D.C.
www.nap.edu

THE NATIONAL ACADEMIES PRESS 500 Fifth Street, NW Washington, DC 20001

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This study was supported by Contract/Grant No. HHSN263000025 between the National Academy of Sciences and the Department of Health and Human Services. Support of the work of the Committee on National Statistics is provided by a consortium of federal agencies through a grant from the National Science Foundation (No. SES-1024012). Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the organizations or agencies that provided support for the project.

International Standard Book Number-13: 978-0-309-30689-0
International Standard Book Number-10: 0-309-30689-2

Additional copies of this report are available from the National Academies Press, 500 Fifth Street, NW, Keck 360, Washington, DC 20001; (800) 624-6242; http://www.nap.edu.

Copyright 2014 by the National Academy of Sciences. All rights reserved.

Printed in the United States of America.

Suggested citation: National Research Council and Institute of Medicine. (2014). *The National Children's Study 2014: An Assessment.* Panel on the Design of the National Children's Study and Implications for the Generalizability of Results, G.J. Duncan, N.J. Kirkendall, and C.F. Citro (Eds). Committee on National Statistics, Division of Behavioral and Social Sciences and Education, and Board on Children, Youth, and Families, Institute of Medicine. Washington, DC: The National Academies Press.

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. C. D. Mote, Jr., is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Victor J. Dzau is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. C. D. Mote, Jr., are chair and vice chair, respectively, of the National Research Council.

www.national-academies.org

PANEL ON THE DESIGN OF THE NATIONAL CHILDREN'S STUDY AND IMPLICATIONS FOR THE GENERALIZABILITY OF RESULTS

Greg J. Duncan (*Chair*), School of Education, University of California, Irvine
Dean B. Baker, Division of Occupational and Environmental Medicine, School of Medicine, University of California, Irvine
Paul P. Biemer, RTI International, Research Triangle Park, Raleigh, NC, and University of North Carolina at Chapel Hill
Barbara Lepidus Carlson, Mathematica Policy Research, Cambridge, MA
Ana V. Diez-Roux, School of Public Health, Drexel University
Virginia M. Lesser, Department of Statistics, Oregon State University
Marie C. McCormick, Department of Social and Behavioral Sciences, Harvard School of Public Health
Sara S. McLanahan, Department of Sociology and Woodrow Wilson School of Public and International Affairs, Princeton University
George R. Saade, University of Texas Medical Branch, Galveston
S. Lynne Stokes, Department of Statistical Sciences, Southern Methodist University
Leonardo Trasande, Departments of Population Health, Environmental Medicine, and Pediatrics, New York University Langone Medical Center

Nancy J. Kirkendall, *Study Director*
Agnes E. Gaskin, *Administrative Assistant*

COMMITTEE ON NATIONAL STATISTICS
2013-2014

Lawrence D. Brown (*Chair*), Department of Statistics, The Wharton School, University of Pennsylvania
John M. Abowd, School of Industrial and Labor Relations, Cornell University
Mary Ellen Bock, Department of Statistics, Purdue University
David Card, Department of Economics, University of California, Berkeley
Alicia Carriquiry, Department of Statistics, Iowa State University
Michael E. Chernew, Department of Health Care Policy, Harvard Medical School
Constantine Gatsonis, Center for Statistical Sciences, Brown University
James S. House, Survey Research Center, Institute for Social Research, University of Michigan
Michael Hout, Department of Sociology, New York University
Sallie Keller, Virginia Bioinformatics Institute at Virginia Tech, Arlington
Lisa Lynch, The Heller School for Social Policy and Management, Brandeis University
Colm O'Muircheartaigh, Harris School of Public Policy Studies, University of Chicago
Ruth Peterson, Criminal Justice Research Center, Ohio State University
Edward H. Shortliffe, Department of Biomedical Informatics, Columbia University, and Department of Biomedical Informatics, Mayo Clinic Campus of Arizona State University
Hal Stern, Donald Bren School of Information and Computer Sciences, University of California, Irvine

Constance F. Citro, *Director*
Jacqueline R. Sovde, *Program Coordinator*

BOARD ON CHILDREN, YOUTH, AND FAMILIES

Angela Diaz *(Chair),* Departments of Pediatrics and Preventive Medicine, Icahn School of Medicine at Mount Sinai, New York
Shari Barkin, Monroe Carell Jr. Children's Hospital, Vanderbilt University
Thomas F. Boat, College of Medicine, University of Cincinnati
W. Thomas Boyce, Faculty of Medicine, University of British Columbia
David A. Brent, Western Psychiatric Institute and University of Pittsburgh School of Medicine
David V.B. Britt, Leadership Development Consultant, Amelia Island, FL
Debbie I. Chang, Nemours Health and Prevention Services, Newark, DE
Janet M. Currie, Woodrow Wilson School of Public and International Affairs, Princeton University
Patrick H. DeLeon, F. Edward Hebert School of Medicine and the Graduate School of Nursing, Uniformed Services University of the Health Sciences, Bethesda, MD
Gary W. Evans, Department of Design and Environmental Analysis and Department of Human Development, Cornell University
Elena Fuentes-Afflick, Department of Pediatrics, University of California, San Francisco, and San Francisco General Hospital
Eugene E. Garcia, Mary Lou Fulton Teachers' College, Arizona State University
J. David Hawkins, School of Social Work, University of Washington
Jeffrey W. Hutchinson, Uniformed Services University of the Health Sciences, Bethesda, MD
Jacqueline Jones, Educational Research Consultant, Princeton, NJ
Ann S. Masten, Institute of Child Development, University of Minnesota
Velma McBride Murry, Peabody College, Vanderbilt University
Bruce S. McEwen, Harold and Margaret Milliken Hatch Laboratory of Neuroendocrinology, The Rockefeller University
Pamela Morris, The Steinhardt School of Culture, Education, and Human Development, New York University
Taha E. Taha, Bloomberg School of Public Health, Johns Hopkins University

Kimber Bogard, *Board Director*

Acknowledgments

The Panel on the Design of the National Children's Study and Implications for the Generalizability of Results thanks the many people and organizations that contributed to the preparation of this report. Without their help, the panel could not have completed its work within the short time frame of the study. As chair, I first thank my fellow panel members for their commitment to the work under a demanding time schedule. They have consistently provided insightful and constructive input under tight deadlines as we developed this report.

The panel thanks Steven Hirschfeld and Jennifer Kwan of the Program Office of the National Children's Study (NCS), at the U.S. Department of Health and Human Services, for their collegial interactions with the panel and their responses to our many questions. The panel also thanks its expert cost consultants, Randall Olsen of Ohio State University and Lisa Schwartz of Mathematica Policy Research, whose input on the cost model and related topics considered by the panel made a significant contribution to our work. The panel also thanks Nicole C. Deziel of Yale University for elaborating on her discussion of environmental measures at the National Research Council/Institute of Medicine Workshop on the Design of the National Children's Study in January 2013.The panel also thanks Kate H. Choi, of Princeton University, for her fertility tabulations.

We very much appreciate the input from and thank the representatives of outside organizations who spoke with the panel at our open meetings: Virginia Delaney-Black, of the Children's Hospital of Michigan, who spoke on behalf of a large number of former NCS Vanguard principal investigators; Jaqueline

Mosby, director of the Office of Children's Health Research of the U.S. Environmental Protection Agency; James Perrin, president of the Academy of Pediatrics; and Cynthia Bearer of the Children's Environmental Health Network.

The panel has been assisted by an able staff: Nancy Kirkendall, our study director, who facilitated panel activities and guided the panel on reaching conclusions and writing its report; Agnes Gaskin, who made all arrangements and managed logistics for the panel; and Connie Citro, director of the Committee on National Statistics, whose insights, detailed knowledge, and wise council were invaluable to the panel in its work.

This report has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the Report Review Committee of the National Research Council. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making its published report as sound as possible and to ensure that the report meets institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process. We wish to thank the following individuals for their review of this report: Cynthia F. Bearer, Department of Pediatrics, University of Maryland Hospital for Children; Barbara D. Boyan, School of Engineering, Virginia Commonwealth University; Manning Feinleib, Bloomberg School of Public Health, Johns Hopkins University; Stephen E. Fienberg, Department of Statistics, Carnegie Mellon University; Lynn R. Goldman, Milken Institute School of Public Health, George Washington University; Heather Joshi, Institute of Education, Centre for Longitudinal Studies, London; Roderick J.A. Little, Department of Biostatistics, School of Public Health, University of Michigan; Colm A. O'Muircheartaigh, Harris School of Public Policy Studies, University of Chicago; Nigel Paneth, Departments of Epidemiology and Biostatistics and Pediatrics and Human Development, College of Human Medicine, Michigan State University; and Thomas S. Weisner, Department of Anthropology, Department of Psychiatry and Biobehavioral Sciences, and Center for Culture and Health, University of California, Los Angeles.

Although the reviewers listed above have provided many constructive comments and suggestions, they were not asked to endorse the conclusions or recommendations nor did they see the final draft of the report before its release. The review of this report was overseen by Janet Currie, Center for Health and Well-Being, Woodrow Wilson School of Public and International Affairs, Princeton University, and Johanna T. Dwyer, School of Medicine, Frances Stern Nutrition Center, Tufts Medical Center, and Office of Dietary Supplements, National Institutes of Health. Appointed by the National Research Council, they were responsible for ensuring that an independent examination of this report was carried out in accordance with institutional procedures and that all

review comments were carefully considered. Responsibility for the final content of this report rests entirely with the authoring committee and the institution.

Finally, we recognize the many federal agencies that support the Committee on National Statistics directly and through a grant from the National Science Foundation. Without their support and their commitment to improving the national statistical system, the panel's work that is the basis of this report would not have been possible.

Greg J. Duncan, *Chair*
Panel on the Design of the National Children's Study and Implications for the Generalizability of Results

Contents

SUMMARY **1**

1 INTRODUCTION **11**
Charge to the Panel, 11
History of the NCS, 14
Proposed Design of the Main Study, 22
This Panel's Approach, 23
Organization of the Report, 25

2 STUDY DESIGN **27**
The Panel's Conceptual Framework, 27
Design Principles, 29
Key Design Features, 30
Supplemental Samples, 37
Health Disparities, 42

3 SAMPLE DESIGN **45**
Proposed Sample Design, 45
Expected Elements in a Sampling Plan, 48
Evaluation of the NCS Sampling Plan, 50
Next Steps for the Sample Design, 67

4 STUDY CONTENT **71**
Study Content Concepts and Parameters, 71
Proposed Data Collection, 78
Health Disparities, 93
Data Release, 94

5 DATA COLLECTION COSTS **99**
Assumptions, 99
Cost Drivers, 100
Field Cost Estimates: Alternative Models, 101
Evaluation, 105

6 STUDY LEADERSHIP AND SCIENTIFIC OVERSIGHT **107**
Areas of Concern, 107
Approaches to Consider, 111
Lessons from Other Studies, 114

7 SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS **115**
Conceptual Framework and Implications for Design, 116
Probability Sample Design, Size, and Composition, 120
Content and Visit Schedule, 123
Field Costs, 126
Leadership of the National Children's Study, 127

REFERENCES: GENERAL **129**

REFERENCES: NATIONAL INSTITUTE OF CHILD HEALTH AND HUMAN DEVELOPMENT **133**

APPENDIXES

A Communications between the Panel and the NCS Program Office **135**
B Field Costs for the National Children's Study: First 7 Years **145**
C Biographical Sketches of Panel Members and Staff **161**

Summary

The National Children's Study (NCS) was authorized by the Children's Health Act of 2000 (Public Law 106-310) and is being implemented by a dedicated Program Office in the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD). In the words of the Program Office, the NCS is planned to be a "longitudinal observational birth cohort study to evaluate the effects of chronic and intermittent exposures on child health and human development in U.S. children." The NCS would be the first study to collect a broad range of environmental exposure measures for a national probability sample of about 100,000 children followed from birth or before birth to age 21 and has great potential value, as summarized in Box 1-2 in Chapter 1.

Detailed plans for the NCS were developed by 2007 and reviewed by an outside panel (see National Research Council and Institute of Medicine, 2008). At that time, sample recruitment for the NCS Main Study was scheduled to begin in 2009 and to be completed within about 5 years. However, results from the initial seven "Vanguard Study" (pilot) locations, which recruited sample cases in 2009-2010, indicated that the proposed household-based recruitment approach would be more costly and time consuming than planned. In response, the Program Office implemented a number of pilot tests in 2011 to evaluate alternative recruitment methods. Based on these results, the study design was revised in early 2013, and a tentative 2015 start date was set for the Main Study.

In March 2013, Congress requested a review of the revised study design by a panel of the National Research Council and Institute of Medicine, stipulating that contracts were not to be let for the NCS Main Study until 60 days after

completion of the review report. Congress specifically stated the panel study would be "to conduct a comprehensive review and issue a report regarding proposed methodologies for the NCS Main Study, including whether such methodologies are likely to produce scientifically sound results that are generalizable to the United States population and appropriate subpopulations."

NICHD specified that the panel's review should cover such aspects of the Main Study design as the national probability sample's overall sample size and design; the use of hospitals and birthing centers as the primary sampling unit; the relative size of the prenatal and birth strata in the probability sample; the size of the supplemental convenience sample; optimal use of sibling births; use of health care providers to refer prospective participants; the proposed study visit schedule, with emphasis on more frequent data collection in pregnancy and early childhood; the proposed approach to assess health and developmental phenotypes; and the proposed approach to define and characterize health disparities.

OVERALL ASSESSMENT

The National Children's Study has the potential to add immeasurably to scientific knowledge about the impact of environmental exposures, broadly defined, on children's health and development in the United States. The panel supports a number of elements of the proposed design for the NCS Main Study, including;

- the use of a national equal probability sample for a large cohort of births,
- the concept of the study as a data collection platform with a focus on health and development guided by exemplar scientific hypotheses,
- the inclusion of siblings born within the 4-year recruitment window, and
- the collection and storage of biological and environmental samples to permit subsequent analysis.

The panel does not endorse two other aspects of the proposed design: (1) the plan to recruit only one-half of the 90,000 births in the probability sample prenatally and the other half at the time of birth and (2) the proposed convenience samples. The scientific consensus on the importance of prenatal exposures on child health and development strongly supports the need to recruit almost all of the births prenatally (excepting only when the mother does not seek prenatal care). The panel's cost analysis shows that it is feasible to have close to 100 percent prenatal recruitment by dropping the planned convenience samples, which add little scientific value.

For other aspects of the proposed design—including the choice of hospi-

tals as primary sampling units instead of geographic areas as in the early pilot testing, the quality of available hospital sampling frames, whether stratification is adequate for understanding health disparities, the details of the sampling and recruitment strategies, the scientific merit of the proposed exemplar hypotheses that are to guide data collection, the schedule and content of data collection in early waves, and the extent of data collection burden on respondents—the panel did not receive sufficiently detailed information from the Program Office to assess them.

Because of this lack of information and related reasons (detailed in Chapter 6), the panel concludes that achieving a scientifically grounded and cost-effective design and implementation for the Main Study will require expansion of the scientific expertise in the Program Office, establishment of an authoritative multidisciplinary oversight structure to review the Program Office's decisions, and regular independent outside review.

STUDY DESIGN

The panel evaluated whether the current NCS plan embraces a scientific approach that identifies current and anticipated future subject domains and measures that are of high priority for understanding children's health and development. Such an approach is needed to guide key study design elements, such as the target population, the sampling strategy, and the schedule and content of data collection.

The panel endorses a number of the general elements of the NCS study design framework that are consistent with best design principles for longitudinal birth cohort studies. These elements include a conception of the study as a data collection platform (noted above) and a dynamic conception of health and disease, which calls for measuring health status, disease conditions, symptoms, and behaviors rather than just existing disease categories. For the latter element, however, insufficient details about the additional measures of conditions and symptoms to be collected precluded the panel's ability to determine whether the burden on respondents could be excessive.

The panel also accepts the NCS assertion that scientific discovery during the more than two-decade life of the study will be facilitated by using "exemplar hypotheses" to guide the study design development rather than enumerating an extensive list of specific hypotheses. At the same time, the panel finds that the exemplar hypotheses proposed for the NCS are not sufficiently well developed to guide sample design and data collection in the early waves, nor is there a clear plan to identify lines of inquiry that could lead to important exemplar hypotheses to guide data collection in later waves.

As noted above, the panel does not endorse one important aspect of the proposed NCS design: the plan to recruit half the cases prenatally and half at birth. This plan does not adequately reflect the growing scientific consensus on

the importance of prenatal influences on child health and development. The panel recommends that the NCS use prenatal recruitment and data collection for almost all sample members, making an exception only for the small percentage of births to mothers who did not receive prenatal care. As part of its evaluation, the panel commissioned a cost analysis to enable it to identify alternative sample designs that could achieve nearly 100 percent prenatal recruitment and be largely cost neutral with the design as currently proposed.

SAMPLE DESIGN

The current NCS design calls for a national, equal probability of selection sample of 90,000 births using a stratified list sample design: hospitals are the primary sampling units, and prenatal care providers whose patients deliver at the selected hospitals are the secondary sampling units. In the proposed design, recruitment would occur during the prenatal period in providers' offices for half of the sample and shortly after birth in the hospitals for the other half of the sample. The proposed sample design also includes 10,000 "convenience sample" births, 5,000 of which would be born to first-time mothers recruited prior to their child's conception. The other 5,000 would be available to address a variety of issues that the NCS would define later: examples include populations exposed to natural disasters, such as hurricanes or industrial accidents, and children likely to experience disparities in health outcomes and not adequately represented in the 90,000-birth main sample.

The panel endorses a number of key aspects of this sample design, including its large size, use of probability-based selection methods, and intent to sample births with approximately equal selection probabilities. Plans for sample stratification (for example, by race and ethnicity) were insufficiently detailed for the panel to evaluate whether they are adequate to respond to the mandate for health disparities research priorities in the Children's Health Act of 2000.

For a variety of reasons (detailed in Chapter 2), the panel judges the added value of the 10,000-birth convenience sample cases to be substantially less than the value of additional cases in the probability sample, recruited prenatally. The panel's cost analysis shows that, holding total field costs constant, the elimination of the planned convenience samples would allow the probability sample to increase in size from 90,000 births, half of which are recruited prenatally, to nearly 95,000 births, nearly all of which are recruited prenatally.

In addition to moving from household-based recruitment to provider-based recruitment, a key change in the NCS sample design since the 2008 review is from a county-based to a hospital-based primary sampling frame. Very few details regarding the hospital-based frame were made available to the panel, in part because of a decision by NCS to change the source of the hospital list in early fall 2013. A list of hospitals and birthing centers might provide a useful frame for the NCS first-stage sample: however, the panel lacked information

concerning the quality of the proposed frame to determine whether it is justified for the NCS to shift from a county- to a hospital-based approach.

Another change from the initial 2008 design is the proposed inclusion of all siblings born subsequently to the originally sampled "target" children and within the planned 4-year enrollment period. The panel endorses this change, as siblings provide many analytic advantages, most prominently collection of preconception exposure information for second and higher-parity births.

Overall, however, the documents that NCS made available to the panel did not provide sufficient details for an evaluation of whether the proposed sample would meet the minimal standards of a scientifically based sample design required for large national data collections. The panel considers it vital for the NCS to develop a detailed sampling plan and recruitment strategy for the Main Study, using survey experts who have extensive experience in conducting large national surveys. The resulting detailed plan should also be appropriately reviewed.

STUDY CONTENT

The panel was asked to consider the proposed frequency of data collection, comparing the early and later years of the study. Although there is reasonable scientific justification to conduct more frequent data collection during the prenatal period and early years, the panel did not receive an adequate explanation of the scientific basis for the specific proposed schedule of visits.

With regard to study measures and methods, scientific quality is enhanced by using the most valid and standardized data collection measures and methods that are feasible, while maintaining sufficient flexibility to assess emerging domains of scientific inquiry. Since the panel did not receive information on specific study protocols, data collection methods, or study instruments, its review could not address the scientific merit or quality of these aspects of the NCS data collection.

The NCS Program Office has an elaborate process for the identification of study domains, data collection protocols, and instruments. As described in the documents, the process would be guided by a conceptual framework for health and development, including a review of existing standardized assessment methods, and would involve consultation with multiple advisory bodies. Although the process is elaborate, the NCS did not provide to the panel any documentation or draft protocols or instruments that would demonstrate its effectiveness as a means to develop scientifically and operationally optimal data collection protocols.

In addressing its goals of understanding the effects of chronic and intermittent environmental exposures on child health and human development, the NCS proposes to devote considerable resources to obtaining biological and environmental samples and storing them in ways that make them available for

future investigations. The panel endorses this method of ensuring that future innovations in measurement are incorporated into the study.

PROJECT MANAGEMENT

Throughout the panel's study, the Program Office staff impressed us with their willingness to provide responses to the panel's myriad questions on a tight time schedule. However, many of those responses did not provide adequate, well-justified scientific answers. The panel concludes that the office lacks sufficient in-house expertise in relevant scientific and survey research disciplines to enable it to use effectively the input it receives from contractors and advisory groups for design and operational decisions for the NCS Main Study. Moreover, the panel concludes that the current management structure—with the Program Office making decisions for the NCS without authoritative oversight that reflects all areas of relevant expertise—is not likely to produce the optimal design for a study that needs to be implemented in a scientifically grounded and cost-effective manner.

Cost-effective and scientifically grounded operation of the NCS Main Study requires a broader and deeper base of scientific expertise for the Program Office; an authoritative multidisciplinary oversight structure to ensure that the decisions of the Program Office are appropriately vetted by all relevant experts; and provision for periodic comprehensive reviews of the study by an independent outside group. The panel's recommendations to address these shortcomings are below.

CONCLUSIONS AND RECOMMENDATIONS

The NCS Main Study offers enormous potential, but it also presents a large number of conceptual, methodological, and administrative challenges. In addition, funding uncertainties make it difficult to plan a study of this magnitude and duration. Like the scientists associated with the study itself, we are eager for it to succeed. We present our recommendations in the hope that, as it goes forward, the NCS will achieve its intended objective to examine the effects of environmental influences on the health and development of American children.

The panel's detailed conclusions can be found in Chapter 7.

RECOMMENDATION 2-1: The scientific framework for the National Children's Study should be based on current understanding of the determinants of children's health and development and an informed consideration of the likely future trajectory of scientific discovery. The paradigms of developmental biology and life-course epidemiology, coupled with findings from other social and behavioral science research on the prenatal and

early life periods, should guide development of the design for the Main Study.

RECOMMENDATION 2-2: In order to facilitate scientific discovery during and after National Children's Study data are gathered, the Main Study should use a national probability sample with the largest feasible sample size and an approximately equal probability of selection design, and it should recruit nearly all of the cohort as early in pregnancy as possible.

RECOMMENDATION 2-3: In order to facilitate scientific discovery during and after National Children's Study (NCS) data are gathered, the Main Study should use valid and standardized data collection measures and methods, while maintaining flexibility to revise or develop new instruments. The NCS should also use state-of-the-art procedures to collect, archive, and provide access to biological and environmental specimens for future analyses.

RECOMMENDATION 2-4: The proposed strategy for the National Children's Study Main Study to collect detailed data on children's health status, conditions, symptoms, and behaviors should be followed to the extent possible, taking into account constraints of costs, operational feasibility, and the need to not overburden respondents.

RECOMMENDATION 2-5: While the panel appreciates the possible scientific value of gathering preconception exposure information on 5,000 first-birth children as part of the National Children's Study Main Study, this supplemental sample should be dropped because of high costs, the lack of any evidence of the value of such a sample, the lack of detailed plans for both selection and analysis, and potential limitations in the proposed data collection schedule.

RECOMMENDATION 2-6: The other supplemental convenience samples proposed for the National Children's Study Main Study should be dropped from the design, including samples of children exposed to natural disasters or geographically defined environmental exposures, samples of additional members of disadvantaged groups, and samples of siblings born outside the 4-year birth window. The potential added value of the supplemental sample cases is less than the value of the additional cases in the probability sample they would replace, specifically, the value of the additional prenatal cases in the probability sample.

RECOMMENDATION 3-1: The National Children's Study Main Study sample should be stratified by characteristics that will achieve variability in socioeconomic status within important population groups to support analysis of health disparities, as well as achieving variability in environmental exposures and geography to support analysis of relationships between exposures and health outcomes.

RECOMMENDATION 3-2: A detailed plan for sampling, recruitment, and minimizing attrition bias for the National Children's Study (NCS) Main Study should be fully developed and evaluated by sampling and survey experts independent from the NCS and approved by the proposed independent oversight committee before the study moves forward.

RECOMMENDATION 4-1: Prior to proceeding with the Main Study, the National Children's Study (NCS) should develop scientifically well-grounded exemplar hypotheses that should be used to guide and evaluate decisions regarding the NCS design and data collection schedule and domains.

RECOMMENDATION 4-2: Because hypotheses will change over time, the National Children's Study should implement a strong and public process to revise and develop new exemplar hypotheses to guide future study implementation, engaging with the extramural and intramural research communities.

RECOMMENDATION 4-3: The National Children's Study Main Study should collect data during the prenatal period at multiple times for as many of the study participants as the budget will allow.

RECOMMENDATION 4-4: Although the panel does not endorse the current proposal for a substantial birth enrollment stratum, if the National Children's Study (NCS) Main Study retains such a stratum, the NCS should conduct a full pilot test of recruitment and data collection during the birth visit before the Main Study is implemented.

RECOMMENDATION 4-5: The National Children's Study Program Office should document and provide justification for development of the data collection schedule, content, and methods now and going forward. The documentation should be sufficient to guide use of the study data by future researchers.

RECOMMENDATION 4-6: The National Children's Study Program Office should finalize the study visit data collection protocols that it

intends to use for the Main Study (including questionnaires and other measurements), at least through age 1, and then pilot test the protocols before implementing the Main Study. The protocols and findings of the pilot tests should be peer reviewed and approved by the proposed independent oversight committee prior to initiating the Main Study.

RECOMMENDATION 4-7: The relevance to health disparities should be an explicit criterion for selecting the constructs that will be assessed as part of the National Children's Study (NCS) Main Study, the measures that will be used to assess them, and the timing of the assessments. The NCS should obtain input from experts on health disparities in childhood as part of the documented process through which the measures for inclusion are selected, and the measures should be approved by the proposed oversight committee.

RECOMMENDATION 4-8: The National Children's Study should consider producing an "early release" version of the data from the Main Study that includes data collected in the early years of each wave's data collection cycle and makes those data available to analysts under the terms of restricted access data centers.

RECOMMENDATION 5-1: Given the goal for the National Children's Study (NCS) to understand the links of environmental exposures to child health and development and its cost structure, if major reductions in the cost of the study need to be made, they should be reductions in sample size rather than exposure domains. Along with such a decision to reduce the sample size, the NCS should reconsider whether to oversample minorities in order to maintain the ability to evaluate health disparities with a reduced sample.

RECOMMENDATION 6-1: The National Institute of Child Health and Human Development (NICHD) should consider and implement one or more means to enhance the scientific expertise of the National Children's Study (NCS) Program Office by recruiting experts in relevant fields from within the National Institutes of Health, other federal agencies, and outside government. In addition, NICHD should consider contracting with experts outside of government to work part time on the NCS as a means to bolster the scientific expertise that is focused on the NCS.

RECOMMENDATION 6-2: The National Institutes of Health should strengthen the oversight and leadership of the National Children's Study (NCS) by establishing an oversight scientific management structure to

include a full range of relevant expertise, with review and approval authority for NCS design decisions.

RECOMMENDATION 6-3: The National Children's Study (NCS) Program Office should establish a mechanism, such as a study section like those in the National Institutes of Health, or use a qualified independent organization to conduct periodic comprehensive outside scientific reviews of the design and operations of the NCS Main Study. To facilitate the work of such a committee and transparency for the study more generally, the NCS Program Office should promptly post on its website all scientific studies conducted for the NCS.

1

Introduction

As authorized by the Children's Health Act of 2000, the National Children's Study (NCS) is planned to be a "longitudinal observational birth cohort study to evaluate the effects of chronic and intermittent exposures on child health and human development in U.S. children . . ." with "a cohort size of approximately 100,000 children and a national probability sample" (NICHD, 2013b, p. 3; see Box 1-1). The NCS will follow children from before birth or at birth to 21 years of age with great promise to contribute to the understanding of the impact of the environment, broadly defined, on the health and well-being of children (see Box 1-2).

The Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) Office of the Director has primary responsibility for planning and coordinating the NCS. This chapter provides the detailed charge to the panel, a brief history of the NCS, a summary of the proposed design of the NCS Main Study, the approach of the panel, and an overview of this report.

CHARGE TO THE PANEL

Section 1508 of the *Consolidated and Further Continuing Appropriations Act, 2013*, states:[1]

> That the Director (of NIH) shall contract with the National Academy of Sciences within 60 days of enactment of this Act to appoint an expert Institute of Medicine/National Research Council (IOM/NRC) panel to

[1]Available: http://thomas.loc.gov/cgi-bin/query/z?c113:H.R.933.enr [March 2014].

conduct a comprehensive review and issue a report regarding proposed methodologies for the NCS Main Study, including whether such methodologies are likely to produce scientifically sound results that are generalizable to the United States population and appropriate subpopulations: *Provided further,* That no contracts shall be awarded for conducting the Main Study until at least 60 days after the IOM/NRC report has been available to the public.

The NICHD requested that the Committee on National Statistics of the National Research Council, in collaboration with the Board on Children, Youth,

BOX 1-1
Children's Health Act of 2000
Public Law 106-310

SEC. 1004. Long-Term Child Development Study

a. *PURPOSE. It is the purpose of this section to authorize the National Institute of Child Health and Human Development (NICHD) to conduct a national longitudinal study of environmental influences (including physical, chemical, biological, and psychosocial) on children's health and development.*
b. *IN GENERAL.—The Director of the National Institute of Child Health and Human Development shall establish a consortium of representatives from appropriate Federal agencies (including the Centers for Disease Control and Prevention, the Environmental Protection Agency) to—*
(1) plan, develop, and implement a prospective cohort study, from birth to adulthood, to evaluate the effects of both chronic and intermittent exposures on child health and human development; and
(2) investigate basic mechanisms of developmental disorders and environmental factors, both risk and protective, that influence health and developmental processes.
c. REQUIREMENT.—The study under subsection (b) shall—
(1) 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;
(2) gather data on environmental influences and outcomes on diverse populations of children, which may include the consideration of prenatal exposures; and
(3) consider health disparities among children, which may include the consideration of prenatal exposures.

NOTE: Congress subsequently passed Public Law 110-154 in 2007 to rename the institute as the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development.

SOURCE: The full text of the law is available at http://www.gpo.gov/fdsys/pkg/PLAW-106publ310/html/PLAW-106publ310.htm [December 2013].

and Families of the National Research Council and the Institute of Medicine, assemble an ad hoc panel to conduct this congressionally mandated review of the design of the NCS Main Study.[2] The NICHD charge to the panel is narrower than the "comprehensive review" language in the congressional act:

> An ad hoc panel will conduct a congressionally mandated review of the design of the National Children's Study (NCS) Main Study. The NCS is intended to follow a cohort of 100,000 children identified at or before birth through age 21 years. The study consists of a pilot or Vanguard Study, currently under way, and a Main Study, which is not to begin until after the expert panel has delivered its report. Based on Vanguard Study experience and other input, as of the meeting of the NCS Federal Advisory Committee on February 26, 2013, the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) proposed a multi-stage probability design for 90 percent of the expected sample size of children for the Main Study. The first stage would select a national probability sample of hospitals and birthing centers from a national listing. A second stage would list prenatal care providers that "feed" patients for delivery at the selected hospital or birthing center, and a third stage would attempt to recruit pregnant women seeing these providers during their prenatal period. In addition, some women may need to be enrolled at the hospital at delivery. The remaining 10 percent of the Main Study sample is set aside for targeted populations to address additional questions of scientific interest, including subsequent births to women selected into the probability sample. The children enrolled in the Main Study will have a wide range of data collected about them, their parents, and their environment at specified intervals over the life of the study.
>
> The charge to the expert panel is to review this proposed design with regard to the sampling frame, the sample design, the recruitment and retention process, and broad aspects of the interview schedule and data collection procedures to determine their scientific merit and, in particular, to determine the expected generalizability of results to a national population and population groups. The panel's review will cover such aspects of the Main Study design as the national probability sample's overall sample size and design; the use of hospitals and birthing centers as the primary sampling unit; relative size of the prenatal and birth strata in the probability sample; the size of the supplemental convenience sample; optimal use of sibling births; use of health care providers to refer prospective participants; proposed study visit schedule, with emphasis on more frequent data collection in pregnancy and early childhood; proposed approach to assess health and developmental phenotypes; and proposed approach to define and characterize health disparities.

[2]The NCS Main Study is distinguished from the pilot "Vanguard Study," which is currently under way.

The panel will deliver a report with conclusions and recommendations at the end of the study that take cognizance of logistical and resource constraints as provided by NICHD.

HISTORY OF THE NCS

The NCS has had a long gestation period, which in part reflects the challenges in addressing some of the issues identified in the panel's charge. Fluctuations in its budget reflect some of this history: see Table 1-1.[3] The table

[3]For a brief history of the NCS up to about 2008, see National Research Council and Institute of Medicine (2008).

BOX 1-2
The Promise of the National Children's Study

As envisioned by Congress, the National Children's Study (NCS) has great potential to address the major effects on and high costs of child morbidity due to potentially preventable conditions in the United States. This potential reflects the growing body of investigation showing that social, economic, and environmental exposures that are encountered prenatally and during early life and that accumulate throughout childhood have long-term consequences for social, emotional, physical, and cognitive well-being in adulthood. "Nearly all domains of later health experience, including cardiovascular disease, various cancers, respiratory disease, cognitive decline, and psychological impairment, have been associated with early-life exposures of one kind or another" (Lawlor et al., 2009, p. 897).

Research on early life factors has often stemmed from the exploration of data from historical birth cohorts (e.g., Barker et al., 1989a, 1989b) and from new kinds of analyses of existing birth cohorts (e.g., Hardy and Kuh 2009; Nybo Andersen et al., 2009). The results from those studies, in turn, have stimulated efforts to revitalize older pregnancy and birth cohorts by collecting new information and to establish new cohorts. Both the number and scope of published studies drawing data from birth cohort studies have increased over the past decade (Lawlor et al., 2009).

As Lawlor et al. (2009, p. 897) note, a number of recent national birth cohorts, including the Danish National Birth Cohort, the Norwegian Mother and Child Cohort, and the U.S. National Children's Study: "have recruited, or plan to recruit, 100,000 parents and children, in order to determine the genetic and life-course influences on childhood health, development and/or common complex diseases in adulthood." A more recent birth cohort study of a similar magnitude and scope is the Japan Environment and Children's Study that started recruiting in 2011. The United Kingdom has a strong tradition of national birth cohort studies going back to the 1940s (see Pearson, 2011).* The new Life Study in the United Kingdom is the latest example: it is planned to include 80,000 births to be recruited in pregnancy

shows that no money was appropriated for implementation of the NCS until 2007. The timeline associated with the evolution of the NCS is summarized in Box 1-3 and is described in more detail below. During the early 2000s, the NICHD convened multiple workshops and commissioned several reports on NCS design options. By 2002, a federally chartered advisory group, the NCS Federal Advisory Committee, was established, and in 2003 the NICHD set up a NCS Program Office to plan and manage the NCS.

The decision that a nationally representative probability sample should be a core element of the study was reached in June 2004 by agreement of the NCS Program Office, the Advisory Committee, and an NCS Sampling Design Workshop Panel. Later that year the NCS, in collaboration with the National Center for Health Statistics, developed the NCS first-stage sample, comprising 110

and born in 2015. As noted above, these studies are similar in purpose and scope to the National Children's Study. In many of these countries, unlike the situation in the United States, there are population registers, school records, administrative records for health care providers, and disease registers that facilitate sample design, data linkage, and analysis and thereby potentially decrease data collection costs. However, the scope and quality of the information available in administrative records is often less than what can be gathered in the survey-based approach of the NCS.

The NCS is planned to be the most comprehensive study of child health and development in the world. The study is intended to enroll a nationally representative sample of nearly 100,000 U.S. babies at or before birth and follow them through age 21, gathering detailed information about their environmental exposures, health conditions, and social, emotional, and cognitive development. As Guttmacher et al. (2013, p. 1873) note:

> Other U.S.-based longitudinal studies provide information about children's health, growth, or development. None of them, however, are as large as the NCS, entail collecting such detailed biologic and environmental data and samples, or include longitudinal phenotyping from before birth through the age of 21. To maximize data interoperability, the NCS is being coordinated with similar studies in other countries (including France, Japan, Britain, and Canada), but none of those studies will examine the same populations or environmental factors as the NCS.

The proposed NCS data collection program has the potential to take advantage of emerging innovations in the biological, social-psychological, and environmental sciences, maximizing the chances for scientific discoveries regarding child health and development. Successfully carried out, the NCS has the potential to enable U.S. scientific leadership in child health-related science for the next several decades.

---

*For information about the UK studies, see http://www.cls.ioe.ac.uk/ [May 2014].

TABLE 1-1 Interagency and Congressional Funding for the National Children's Study (dollars in millions)

Fiscal Year	Nominal Value[a]	Present Value (2012)[b]
	Funding Provided by NICHD During the Planning Phase	
2000	1.0	1.4
2001	3.2	4.5
2002	6.1	8.2
2003	10.6	13.6
2004	10.6	13.0
2005	11.0	13.0
2006	12.2	14.0
	Funding Appropriated by Congress During the Implementation Phase	
2007	68.8	76.2
2008	119.9	119.5
2009	179.8	191.9
2010	193.9	200.2
2011	191.1	192.1
2012	193.1	193.1
2013	Up to 165.0	Up to 162.6
2014	Up to 165.0	Not available

[a]Nominal value budget information for 2000 through 2013 taken from http://www.nationalchildrensstudy.gov/about/funding/Pages/interagencycongressionalfunding.aspx [April 2014]; 2014 budget information from 2015 appropriations language from http://officeofbudget.od.nih.gov/pdfs/FY15/Appropriation%20Language.pdf [April 2014].

[b]The present value is computed using the price index of the Bureau of Economic Analysis for government nondefense consumption expenditures dated January 2, 2014; see http://www.bea.gov/national/nipaweb/SS_Data/Section3All_xls.xls [April 2014].

primary sampling units in 105 locations around the country. Beginning in 2005, the NCS Program Office awarded a number of contracts to design and implement the study. The initial NSC research plan was completed in June 2007, and the NICHD requested an external review of the plan, which is the subject of the National Research Council and Institute of Medicine (2008) publication.

The 2007 Study Design

The early architects of the NCS designed a sampling strategy to deliver a national probability sample of 100,000 children using a stratified area sample design that was expected to provide preconception environmental exposure information for about 25 percent of these children and early prenatal information for about 90 percent (see details in Chapter 3). The area sample of 110 primary sampling units (counties, groups of counties, or parts of counties) was designed so that each unit would experience a minimum of 2,000 births during a 4-year birth enrollment period, which would yield a target of 1,000 births for

BOX 1-3
Timeline of the National Children's Study

2000	National Children's Study (NCS) authorized by Congress
2002	Federal Advisory Committee for NCS established
2003	NCS Program Office established within National Institute of Child Health and Human Development
2004	Nationally representative sample determined to be core element of study design
2005	First-stage sample (counties or groups of counties) drawn by National Center for Health Statistics
2007	1. Appropriation for NCS authorized by Congress 2. NCS research plan completed
2008	National Research Council/Institute of Medicine review of NCS research plan published
2009	1. Field operations begun in 7 initial Vanguard locations 2. Change in NCS leadership 3. NCS decision made to expand pilot study to 30 additional locations using field contractors that had already been selected (but not activated) to implement the Main Study in those locations. All 37 locations become part of Vanguard Study
2010	Alternative recruitment strategies tests begin fieldwork (10 locations each for 3 alternatives)
2012	1. Provider-based sampling pilot initiated in 3 additional locations, bringing the Vanguard study to a total of 40 locations to test 5 different recruitment strategies 2. Workshop on sampling strategies sponsored by NCS 3. Workshop on environmental exposures and measures sponsored by the Environmental Protection Agency, National Institute of Environmental Health Sciences, and NCS
2013	1. National Research Council/Institute of Medicine Workshop on sample design conducted and summary issued 2. National Research Council/Institute of Medicine Panel on the Design of the National Children's Study and Implications for the Generalizability of Results established
2014	Publication of National Research Council/Institute of Medicine Panel's report

each unit over the birth period with each birth having an equal probability of selection. Secondary sampling units (called segments) were defined as aggregations of contiguous census blocks that, when summed together within each primary sampling unit, would yield the desired target for the primary sampling unit. The secondary sampling only occurred in the primary sampling units that were designated to participate in one of the recruitment pilot (or Vanguard) studies (see below). The "measure of size" for the segments was estimated from the number of births in the census blocks (geocoded to the mother's place of residence) during a prior 5-year period. Segments were randomly sampled, although in most of the sampling units, contiguous segments were clustered into geographical strata with one segment sampled per stratum in order to enhance the likelihood that the random sample of segments would be diverse and geographically distributed across the primary sampling units.

The original design of the study was organized around hypotheses in 28 topic areas, referred to as core or meta-hypotheses. These hypotheses focused on 7 priority outcome areas:

1. pregnancy,
2. neurodevelopment and behavioral,
3. child health and development,
4. asthma,
5. obesity and growth,
6. injury, and
7. reproductive development.

Each topic had several specific hypotheses with power calculations to demonstrate that the hypotheses could be addressed with the anticipated size of the NCS study population, and specific items of data collection were linked to these topics.

Implementation and Testing

As indicated above, the NCS Program Office was established in 2003 with a few staff and no direct funding. The initial direct appropriation for the NCS was to NICHD, and after 2009 the appropriation for the NCS was to the Office of the Director of the National Institutes of Health. According to the NICHD website, "the NICHD Office of the Director has primary responsibility for planning and coordinating the National Children's Study."[4] The NCS Program Office directs the implementation of the day-to-day operation of the study and houses scientists, staff on detail from other organizations, program analysts, and

[4]Available: https://www.nichd.nih.gov/about/org/od/ncs/Pages/index.aspx [March 2014].

contractors.[5] In 2014, the NCS Program Office had 18 staff members listed on the NCS website. Of them, 4 held medical degrees, 9 held advanced degrees equivalent to a Ph.D.,[6] and the director had both an M.D. and a Ph.D. Six of the 18 were contracting officer representatives.

The NCS Program Office awarded contracts in 2005 to academic programs in seven of the 110 sampled primary sampling units to pilot test the household recruitment and data collection protocols. The NCS uses the term "Vanguard Study" to refer to all pilot-testing activities used to plan the NCS Main Study.[7] The seven units used for the initial pilot testing were called the initial Vanguard locations, and the academic programs that were contracted to implement the field work were called the initial Vanguard centers. Field operations began in 2009 in the seven initial Vanguard locations[8] using a household enumeration and screening strategy to identify eligible women for recruitment. The goal was to enroll approximately 1,750 pregnant women through these seven study locations (250 per location) during 12 months of data collection. The enrolled families were considered to be a "Vanguard cohort" that would be followed for 21 years and used for pilot testing future data collection protocols. The plan was for the NCS Main Study to begin after the 12 months of enrollment in the initial Vanguard locations, with the field work and participant enrollment beginning in three annual waves of approximately one-third of the 110 sampled primary sampling units. The initial Vanguard centers would participate in the Main Study as part of wave 1, while continuing to follow the Vanguard cohort participants.

In early summer 2009, the NCS Program Office indicated concern that the initial enrollment in the Vanguard Study was lower than anticipated. Although the response rates at various stages of recruitment were comparable to or higher than those for other large studies,[9] the pilot study yielded fewer pregnant

[5]Available: http://www.nationalchildrensstudy.gov/about/organization/programoffice/Pages/default.aspx [March 2014]. Includes the names of 18 key staff members, many with a linkage to a brief bio.

[6]Fields of study of staff with a Ph.D. included cell biology, physiological molecular biology, environmental health science, sociology, epidemiology, and health policy and management with a concentration in bioethics. For the other three individuals with a Ph.D. or equivalent, no additional information was available.

[7]The purpose of the Vanguard Study is to pilot test data collection protocols and methods, participant follow-up, personnel training, etc. The panel understands and data users should expect that there will be some changes as aspects of the study design are tested and revised. For this reason and others, data collected in the Vanguard Study will often not be comparable with data collected in the Main Study.

[8]Recruitment began in two locations in February 2009 and in the remaining five locations in April 2009. The implementation was phased across segments (the secondary sampling units), so recruitment did not begin in some segments until November 2009.

[9]For details, see http://www.nationalchildrensstudy.gov/about/organization/advisorycommittee/Pages/ParkC-NCSAC-Jan2010-Recruitment-final.pdf [March 2014, p. 12].

women and births than had been anticipated.[10] Based on this information, "the NCS Program Office projected that enrollment using a household recruitment approach could take much longer to complete and cost more than anticipated" (NICHD, 2013d, p.5). In the summer of 2009, the NCS leadership was changed (for details see Wadman, 2009, pp. 20-21). Among other issues related to this change, it had been reported that the revised projected cost of the study was significantly more than previously planned, even without accounting for new recruitment results. In the fall of 2009, the new NCS leadership decided to conduct additional recruitment pilot studies rather than to proceed immediately to the Main Study.

The field study centers contracted to implement the Main Study[11] became the platform for testing the additional recruitment alternatives, and the funding that was appropriated for the Main Study instead supported the Vanguard Study. This phase of pilot testing was called the alternative recruitment pilot of the Vanguard Study. To test recruitment strategies, the NCS Program Office developed three alternative recruitment strategies and implemented each of them in 10 of the primary sampling units that had been sampled for the Main Study. The locations were selected to be geographically distributed throughout the country and have different population sizes and characteristics, although they were not formally considered to be a representative or probability subsample of the NCS Main Study primary sampling units. The three alternative recruitment strategies were an enhanced household enumeration strategy, a direct outreach recruitment strategy, and a provider-based recruitment strategy (NICHD, 2013d, p. 5). Although the study populations were recruited using different methods, the segment-based secondary sampling units were still used to define residential eligibility.

Evaluation of the three strategies indicated the provider-based recruitment to be the most efficient method of recruiting pregnant women. However, even that approach was inefficient in more populous locations because many different groups were providing prenatal care to women living in the same segment. Since sample eligibility was determined from the mother's location of residence at the time of delivery, many different provider groups had to be engaged. Therefore, the NCS began planning for a new provider-based sampling pilot study that eliminated the geographic secondary sampling units and instead defined the prenatal care providers as secondary sampling units (NICHD, 2013d, p.6). At the end of 2013, with the provider-based sampling study still

[10]By the conclusion of the initial Vanguard pilot in April 2010, only about 800 pregnant women had been enrolled or about 45 percent of the goal. For details, see http://www.nationalchildrensstudy.gov/about/organization/advisorycommittee/Pages/A1_NCS_Alt_Recruitment_Schema.pdf [March 2014, p. 4].

[11]In 2012, the NCS Program Office contracted with four centers, called Regional Operating Centers, to continue the Vanguard Study and to follow families that were previously enrolled. The contracts with the original 40 Vanguard Study centers were subsequently discontinued.

under way, preliminary results indicated that it was an efficient secondary sampling strategy. A challenge that was identified with all of the recruitment methods based on a geographic primary sampling unit, including the provider-based sampling method, was that a large number of hospitals had to be engaged in order to collect data and biological specimens at the time of delivery.

The NCS Program Office sought input from the NCS advisory committee and stakeholders on an appropriate sampling design for the Main Study during 2012. The NCS Program Office prepared materials (NICHD, 2012) to provide background and discussion of alternative approaches including probability samples, convenience samples, and hybrid approaches. Sample design was also the topic of a workshop held in May 2012.[12] In its budget request for 2013, prepared in May 2012,[13] NCS stated

> In assessing alternative sampling strategies, National Children's Study and NIH leadership considered the overall scientific goals and which of these could be achieved with different strategies. Additional considerations were costs, based on Vanguard data, and the reality of flat or shrinking budgets for biomedical research. As a consequence, NIH now proposes that the Main Study sampling frame be based on provider location. One approach for developing such a sampling frame would be to use providers associated with specific health plans. Such an approach would have several advantages in terms of cost and feasibility, but would abandon the geographic based probability sample. Consequently, the enrolled population would no longer be a national probability sample but, instead, a well described cohort followed longitudinally.

First New Sample Design Proposed in 2013

In 2013, NICHD requested that the National Research Council and Institute of Medicine convene a workshop for a public discussion of a proposed new sampling plan for the NCS Main Study, which differed from the approach suggested in its budget request for 2013. In this plan, hospitals identified as secondary sampling units would be selected from a geographically selected primary sampling unit. Prenatal care providers whose patients delivered at selected hospitals (or birthing centers) would serve as third-stage sampling units. A sample of women who visit one of the sampled prenatal care providers and expect to deliver at one of the selected birthing centers or hospitals would be recruited.

One new feature of this plan was that the sample of 100,000 births would be split into a probability sample of 90,000 with two enrollment strata and an

[12]See summary of workshop discussion at http://www.nationalchildrensstudy.gov/research/workshops/Pages/potential-sampling-strategies-minutes-may-2012.pdf [May 2014].

[13]See http://www.nationalchildrensstudy.gov/about/funding/Pages/congressionaljustification.aspx [April 2014].

unspecified convenience sample of 10,000. For the probability sample, 45,000 children would be sampled from prenatal care providers using the selected hospitals, and 45,000 children would be sampled at birth from the same selected hospitals. The NCS explained this allocation in practical terms: "NCS would need more time and resources to enroll pregnant women from prenatal care providers than from birth facilities" (NICHD, 2013b, p. 25). The allocation of the sample between the prenatal care providers and hospitals was discussed at the workshop; the workshop discussions are summarized in National Research Council and Institute of Medicine (2013).

PROPOSED DESIGN OF THE MAIN STUDY

The sampling plan provided to the panel by the NCS Program Office in August 2013 for this study (NICHD, 2013b) was yet further changed from previous plans. This plan relied on a list frame of hospitals from which to select primary sampling units, replacing the geographic-based area sample that was previously proposed. Prenatal providers whose patients deliver at the selected birthing hospitals and birthing centers would serve as secondary sampling units. The current proposal retains the 45-45-10 distribution of the sample among prenatal care providers, birthing hospitals, and the convenience sample, respectively.

At the time the panel was drafting this report, the specific methodologies of the sampling plan were still under development, and many important details were not available to the panel. For example, between August and October of 2013, the frame under consideration switched from a list of hospitals available from the American Hospital Association augmented with a list of birthing centers to a list of hospitals assembled from the 2010 State Inpatient Database maintained by the Agency for Healthcare Research and Quality of the U.S. Department of Health and Human Services.

Several NCS features other than the sampling frame changed since the 2008 review. The current proposal conceptualizes the study less as a vehicle for testing current hypotheses and more as a platform for future researchers to develop and investigate new hypotheses using the previously completed study instruments and archived biological and environmental specimens. Consequently, the current proposal no longer relies on specific hypotheses to define the content of the study. Instead, "the proposed plan was developed using several exemplar hypotheses so it is hypothesis informed but not hypothesis limited" (NICHD, 2013b, p. 34; Appendix 2). The NCS Program Office initially illustrated the impact of sample size by presenting tables using the same hypotheses and same computer code for calculating odds ratios that were used for the 2007 design (p. 36). Later, NICHD listed five exemplar hypotheses (see NICHD, 2013d). The concept of using exemplar hypotheses is discussed in Chapter 2, and the implementation of this strategy is discussed in Chapter 4.

The NCS is also now described as a platform for researchers covering a broad conception of health domains (NICHD, 2013b, p. 36). The NCS will not focus on classifying participants into predetermined disease categories, but instead will collect a set of primary observations and events to enable researchers to apply their own health criteria and form "cases" (p. 31).

Although the use of domains and primary observations, rather than predetermined disease categories, does not necessarily require a shift from the original disease-oriented outcomes of the NCS, the current proposed plan tends to deemphasize a focus on disease outcomes and gives greater emphasis to positive health and development domains. It provides a detailed conceptualization of health and development in seven domains:

1. demographics,
2. physical health,
3. psychosocial health,
4. neurodevelopmental health,
5. health behaviors,
6. social environment, and
7. physical environment.

Each of these domains has subdomains. The NCS will use this conceptualization of health as multidimensional and dynamic to guide the selection of assessments.

With the current plan, NCS proposes to use phenotypes and profiles to describe participants. The NCS explains (NICHD, 2013b, p. 30):

> The term phenotype is used for the observable characteristics including morphology, physiology, developmental stage, behavior and products of behavior. . . . The term profile is used for the larger concept of phenotype plus environmental context. A profile includes observable characteristics about the participant plus information about the environment such as air particle measures, noise level, family structure and dynamics, access to health care, etc.

Thus, at each assessment a participant will be assessed using a health phenotype framework, as well as through collection of environmental data and biospecimens. The NCS will also directly collect some health condition information on all participants.

THIS PANEL'S APPROACH

The panel held three meetings to review the design of the NCS, two of which included 2-hour periods that were open to the public. During the open meetings, NCS staff were present, as were representatives of a group of NCS

Vanguard Study principal investigators, the American Academy of Pediatrics, the Children's Environmental Health Network, and the Environmental Protection Agency's Office of Children's Health. In addition, the panel engaged two consultants, Randall Olsen of Ohio State University and Lisa Schwartz of Mathematica Policy Research, to provide the expert cost analysis included in this report.

In preparation for the panel's work, the NCS Program Office provided a background document to the panel (NICHD, 2013b) and made available background documents it prepared for meetings of its NCS Federal Advisory Committee (NICHD, 2013a, 2013c, 2014b). The volume of background material provided to the panel was considerably less than that provided in the previous external review (see National Research Council and Institute of Medicine, 2008). The earlier study was asked to review a comprehensive research plan that was 745 pages long, including multiple appendixes. For this study, the NCS Program Office supplied a 56-page document that provided an overview of the conceptual framework and processes that would result in the final design of the NCS Main Study.

The panel prepared several sets of questions that were sent to NCS for response. Panel questions and NCS responses were in writing, sent and received by email between the NRC study director and the panel's official point of contact at NCS or the director of NCS. The documents (NICHD, 2013d, 2013e, 2013f, 2013g, 2013h, and 2013j) contain the NCS responses to panel questions. The document (NICHD, 2013i), also provided at the panel's request, is a summary of an interim report concerning the frame and sample of hospitals prepared by NCS consultants. The document (NICHD, 2014a) provides NCS comments on the panel's cost analysis assumptions. Appendix A summarizes the information provided by the NCS Program Office and the questions asked by the panel for which information was not provided. The panel members also reviewed publicly available documents on the NCS Website and searched the peer-review literature for publications related to the NCS and other large national birth cohort studies.[14]

The panel concluded that to meet its charge of evaluating the "scientific merit" of the Main Study, the NCS Program Office would need to provide specific documentation on the sampling design, the sample frame, the data collection protocols, and the study instruments, rather than just describe the conceptual framework, strategies, and anticipated processes to design the Main Study. For example, one key to assessing expected generalizability of results would be detailed information about the proposed sampling plan based on a list of hospitals: the completeness of the list, the number of primary sampling

[14]The separate reference list for NICHD documents includes both those that are publicly available and those that were provided directly to the panel and are available in the Public Access File of the National Academy of Sciences.

units to be selected, stratification, and so on. The panel did not receive such information about the primary sampling units in the current design; the NCS Program Office indicated that the sampling plan was still being developed.

Similarly, the panel was asked to comment on the proposed study visit schedule; the proposed approach to assess health and developmental phenotypes; and the proposed approach to define and characterize health disparities. With these tasks specified, and given that the Vanguard Study had been in place for more than 4 years and that the Main Study was scheduled to begin soon,[15] the panel expected to receive draft final protocols and data collection instruments to assess. However, the panel did not receive these protocols and instruments, apparently because they are still being developed (NICHD, 2013c, 2014b).

Although there was substantial interaction between the panel and the NCS Program Office leaders and staff, the panel could not determine whether all requested documentation was not provided primarily because the study design, protocols, and instruments are incomplete and could not be provided or because there was a difference in perspective between the panel and the NCS Program Office regarding the scope of the panel's charge. Specifically, it was never clear whether the NCS Program Office thought that the panel could evaluate the scientific merit and generalizability of the Main Study by reviewing only the concepts, strategies, and proposed processes rather than reviewing actual documentation of the final proposed study design, protocols, and initial instruments. Regardless of the underlying reason, the documentation the panel requested was not provided: thus, the panel concluded that its review could not address the scientific merit or quality of these critical aspects of the NCS.

ORGANIZATION OF THE REPORT

The remainder of the panel's report includes six chapters, two reference lists, and three appendixes. Chapter 2 describes and comments on overarching features of the current NCS study design, including the conceptual framework and its evolution from a hypothesis-driven and disease-specific approach to a data collection platform with a focus on health. It also describes key issues that drive the sample design, including the utility of the proposed convenience samples. The chapter addresses the following key items the panel was asked to consider: the national probability sample's overall sample size and design; the relative size of the prenatal and birth strata in the probability sample; the

[15]The NCS Program Office stated that they "anticipate that the Main Study will start a few months after the Institute of Medicine/National Research Council report is published." See http://www.nationalchildrensstudy.gov/newsandevents/announcements/Pages/ncsreceivesFY2013appropriationapril2013.aspx [March 2014].

proposed approach to assess health and developmental phenotypes; and the size of the supplemental convenience sample.

Chapter 3 describes and evaluates the proposed sample design in light of the study's history and resource constraints, with more detail on topics introduced in Chapter 2. It comments on statistical issues related to the survey design, as well as on the treatment of special populations. The chapter addresses the following key items the panel was asked to consider: the national probability sample's overall sample size and design; the use of hospitals and birthing centers as the primary sampling unit; the relative size of the prenatal and birth strata in the probability sample; the optimal use of sibling births; and the use of health care providers to refer prospective participants.

Chapter 4 addresses the proposed study visit schedule, the content of collections of data and samples, and the approach to address health disparities. It also reviews the proposed approach to the dissemination of information once the study begins. The chapter comments on the following key items from the panel's charge: the proposed study visit schedule, with emphasis on more frequent data collection in pregnancy and early childhood; and the proposed approach to define and characterize health disparities.

Chapter 5 summarizes the panel's analysis of field costs associated with various alternative sample composition assumptions.

Chapter 6 comments on the study's scientific leadership and need for independent study oversight and periodic outside review.

Chapter 7 summarizes the panel's conclusions and recommendations.

The general reference list includes all documents cited in the report except those from NICHD. The NICHD reference list includes the documents from the agency cited in the report, both those that are publicly available and those that are available in the National Academy of Science's Public Access File.

Appendix A documents the information requested by the panel and information provided by NCS. Appendix B describes the scope and assumptions underlying the panel's field cost analysis in Chapter 5. Appendix C provides brief biographies of the panel members and staff.

2

Study Design

This chapter describes and critiques the proposed study design for the National Children's Study (NCS), including the conceptual framework and its evolution from a hypothesis-driven and disease-specific approach to a data collection platform with a focus on health and development. It also describes overarching issues that drive the sample design, including two of the key topics the panel was asked to consider: the national probability sample's overall sample size and design and the relative size of the prenatal and birth strata in the probability sample.

The chapter also provides the panel's analysis of a third key topic the panel was asked to consider: the proposed uses of supplemental convenience samples to enroll nulliparous women for preconception data collection and to enroll additional populations to address targeted research questions. In this case, we considered the potential value of studying these populations against the overall cost, size, and scope of the probability sample.

Finally, this and subsequent chapters discuss a fourth key topic: strategies for the NCS to address health disparities in children effectively, a charge in the Children's Health Act of 2000.

THE PANEL'S CONCEPTUAL FRAMEWORK

Although nationally representative birth cohort studies have a more than 50-year history, any study like the NCS that aspires to gather exposure and health data over a 21-year period faces difficult design decisions. Many of those decisions have to account for the rapidly changing nature of the relevant

sciences. New environmental dangers are discovered, as are new methods for assaying biological samples and characterizing childhood health conditions. Some of the hypotheses thought to be of greatest interest at the beginning of a study may bear little resemblance to hypotheses that emerge over the course of the study's span of more than two decades. Furthermore, given an ambitious sample size and the likely billions of dollars of total cost, possible budget constraints are also an important concern for the NCS.

Architects of the current study plan for the NCS (Guttmacher et al., 2013) envision it as an ongoing data platform that will support a broad range of scientific discoveries related to the determinants of child health, growth, and development. Other scientific endeavors have provided such data platforms, with recent examples in genetics (the Human Genome Project), astronomy (the Hubble telescope), and particle physics (the Hadron collider). In the social and behavioral sciences, the content and open data policies of long-running national health studies, such as the National Longitudinal Study of Adolescent Health[1] and the Health and Retirement Study,[2] have facilitated discoveries by many social and behavioral scientists and clinicians.

A series of publicly available national-level birth cohort studies, beginning with the 1958 National Child Development Study in Britain, have also supported hundreds of research studies (e.g., Lawlor et al., 2009; Vrjheid et al., 2012). Design lessons for the NCS are also provided by birth cohort studies in the United States, such as the Fragile Families and Child Wellbeing Study[3] and the Department of Education's Early Childhood Longitudinal Study birth cohort sample,[4] as well as studies based on "convenience" rather than representative samples of birth cohorts and longitudinal studies in other countries (see Golding, 2008).

In considering its assigned topics regarding the NCS study design, the panel found it useful to draw from many of these ongoing studies to delineate design principles and guidelines that would optimize the scientific value of a longitudinal birth cohort study of child health and development such as the NCS (e.g., Golding, 2008; Olsen, 2012). Many of these principles mirror key elements of the current NCS study design (detailed in Guttmacher et al., 2013). This chapter lists these design principles and discusses their implications for the NCS study design. It then discusses design features, supplemental samples, and health disparities.

[1]For a description, see http://www.cpc.unc.edu/projects/addhealth [May 2014].
[2]For a description, see http://hrsonline.isr.umich.edu/ [May 2014].
[3]For a description, see http://www.fragilefamilies.princeton.edu/about.asp [March 2014].
[4]For a description, see http://nces.ed.gov/ecls/ [March 2014].

DESIGN PRINCIPLES

We begin our discussion with the principles that bear directly on key elements of the design, sample, and content of the NCS:

- A scientific framework that encompasses current and anticipates future domains of high-priority scientific inquiry is needed to guide key study design elements, such as the target population, the sampling strategy, and the schedule and content of data collection.
- Scientifically robust exemplar hypotheses are needed to guide sample design and early-wave data collection, while decisions about data to be collected in later waves should leave room to take into account new hypotheses that emerge over the course of the study.
- A probability sample ensures that results generalize to the population from which the sample is drawn.
- A large stratified national sample in which all children have an approximately equal chance of selection supports multiple goals. For the NCS, these include estimating relationships between exposures and health outcomes, analyzing health disparities, and attaining representation of children in key demographic and geographic subgroups roughly in proportion to their representation in the population.
- As large a sample size as possible within budget constraints is needed to provide statistical power for current and future scientific discoveries.
- Scientific quality is enhanced by using the most valid and standardized data collection methods that are feasible, while maintaining sufficient flexibility to assess emerging domains of scientific inquiry.
- The study design needs to be as cost effective and as efficient for its key purposes as possible.
- Scientific discovery is enhanced when the potential for future innovations in measurement is incorporated into the study. In the case of the NCS, this argues for collecting and storing biological and environmental samples in ways that make them available for future investigations.
- Discoveries related to health conditions are facilitated by a dynamic conception of health and disease, which calls for measuring health status, disease conditions, symptoms, and behaviors rather than just existing disease categories.
- Discovery is facilitated if data are released as early and as completely as possible, with due regard for the protection of confidentiality.
- Transdisciplinary discovery and statistical sophistication are enhanced when all relevant scientific expertise is integrated into the project management structure.

KEY DESIGN FEATURES

Key design features can be derived from the principles above. This section covers the first nine above; the last two, on data release and project management, are covered in Chapters 4 and 6, respectively.

Scientific Framework

Current understanding of the determinants of children's health and development (e.g., National Research Council and Institute of Medicine, 2000, 2004) and an informed consideration of the likely future trajectory of scientific discovery need to underlie the study. A relatively strong consensus has developed during the past decade about key scientific issues in the field (e.g., Cohen Hubal et al., 2013; Landrigan and Miodovnik, 2011). First, child health, growth, and development is a product of biological factors and a diverse set of environmental influences, including intrauterine and social influences, which implies that high-quality measures of multiple dimensions of both sets of influences need to be taken during appropriate developmental periods.

Second, emerging research on the early origins of future health points to the importance of early postnatal, prenatal, and even preconception conditions, which implies concentrating data collection in the early years relative to the later years of the study. These paradigms of developmental biology and life-course epidemiology, coupled with insights from a number of social and behavioral sciences, should guide development of the NCS study design.

Public health goals are an essential first step in framing hypotheses for a major study, and while public health goals are not explicitly stated by the NCS, according to documentation received by the panel, the NCS is informed by a public health perspective. For example, the approach to identifying the specific domains of health and development (NICHD, 2013b, pp. 29-31, 2013d, pp. 26-29) identify a large number of highly relevant developmental stages, symptoms, and conditions that would include most problems of public health significance. Second, there is a long list of environmental exposures, chemical and socioeconomic, that, if associated with child health and developmental problems, would clearly signal potential for prevention. In addition, public health significance is the first criterion by which data collection will be prioritized (NICHD, 2013d, p. 33, 46-47). Finally, as NCS notes, if realized, the phenotype/health profile approach would facilitate examining conditions over time when coding and diagnostic practices may change (e.g., asking whether a child has autism next year would give a different prevalence than 3 years ago with the shift from the DSM-IV to DSM-V), but using a set of symptoms and treatments, one could define the condition according to whatever diagnostic rubric was being used.

Another important context for the scientific framework is recognition of

the stark nature of persistent differences in health between disadvantaged and more advantaged groups. This issue was acknowledged by Congress, which mandated in the 2000 Children's Health Act that the NCS be designed to "consider health disparities among children." This issue is addressed in more detail below and in Chapter 3.

RECOMMENDATION 2-1: The scientific framework for the National Children's Study should be based on current understanding of the determinants of children's health and development and an informed consideration of the likely future trajectory of scientific discovery. The paradigms of developmental biology and life-course epidemiology, coupled with findings from other social and behavioral sciences research on the prenatal and early life periods, should guide development of the design for the Main Study.

Exemplar Hypotheses

The process of using a scientific framework and context to guide decisions regarding the NCS study design, sampling frame, and data collection protocols can be facilitated by identification of scientifically robust exemplar hypotheses. In the case of the NCS, this means hypotheses that encompass current and anticipate future scientific inquiry concerning high-priority environmental factors and child health and development outcomes, while accounting for potential confounding and effect modification. It also means a need to assess nonpersistent environmental exposures that could occur during periods of developmental plasticity and vulnerability with effects manifested at later stages of development. The NCS Program Office provided exemplar hypotheses in some documents (e.g., NICHD, 2013d, pp. 45-46).

Although some in the scientific community have argued that hypotheses must be specified in advance of undertaking a research project to ensure scientific integrity (e.g., Paneth, 2013), the panel recognizes that it is not possible to anticipate all possible scientific hypotheses that could be addressed over the time span of a long-term study. Yet exemplar hypotheses are critical for guiding sample design and early waves of data collection, while later waves of data collection can also be guided by hypotheses that emerge over the course of the study. Using exemplar hypotheses to guide study design development rather than attempting to develop an exhaustive list of detailed hypotheses is also consistent with the concept of the NCS as a study platform insofar as the study should not be designed only to address specific hypotheses of current interest. Although the number of exemplar hypotheses does not have to be extensive, each must be scientifically robust in order to guide development of the study design. (These issues are addressed in more detail in Chapter 4.)

National Equal Probability of Selection Design and a Stratified Sample

Optimal sample designs are a product of a host of potentially conflicting priorities (Michael and O'Muircheartaigh, 2008). The objectives to study many universal in-the-body biological processes may place minimal constraints on sample selection because the mechanisms may be essentially common to all humans. In contrast, estimating associations between children's health and their cumulative exposures to various physical or social conditions requires samples that provide considerable variation in and covariation among the outcomes and exposures of interest. Furthermore, evaluating health disparities across various population subgroups and describing the prevalence of exposures and health conditions in the population argue for a probability sample that can be statistically weighted to represent the population from which it is drawn. This goal led the NCS Program Office to opt for a national probability sample in the current design, an element that received strong endorsement from the previous study review (National Research Council and Institute of Medicine, 2008) and which we also endorse.

A key design dimension of probability samples is whether all individuals in the population are sampled at similar rates, which is also part of the current NCS design. A large stratified equal probability sample can help ensure that important population subgroups will be represented in the sample in roughly the same proportion as they are in the general population. This dimension is applicable for any major subgroup considered as a stratum in the sampling design. An advantage of this design is that it produces nearly optimal precision for estimating population means or proportions for the population as a whole, and is adequate for important population subgroups, at least at the start of data collection. In addition, because the premise of the currently proposed design is that the population subgroups of interest for the NCS are unknown and cannot be predicted, the fall-back position is to avoid oversampling any one subgroup to minimize the potential harm by necessarily reducing the sample size for some other subgroup. The analysis of new exposures and subpopulations of interest that arise over the course of the study requires a design that provides variation in both known and not-yet-discovered exposures, which is most likely with the current design's proposed equal probability of selection national probability sample.

A possible rationale for unequal selection probabilities arises from the NCS's charge to address health disparities, which by definition involve disadvantaged population subgroups. But, as discussed below and in Chapter 3, we believe that the large size of the main NCS sample, along with careful stratification, will provide sufficient statistical power to investigate health disparities

across the major domains and categories of interest (such as race and ethnicity and socioeconomic status).[5]

Size of the Sample

Determining the optimal sample size is one of the most critical decisions for any large scale epidemiological or population study. As noted above, this decision should be made in the context of the scientific framework and facilitated by the use of exemplar hypotheses that can be used to develop estimates of minimum detectable effect sizes and statistical power. The current and anticipated future scientific issues related to children's health and development to be addressed by the NCS involve complex etiologies with covariation and interaction between multiple factors that can vary over time. Some exposures and health outcomes of great importance may be relatively rare, while others may be common but complex. In order to address the range and complexity of issues that potentially should be addressed by the NCS, the study sample size should be as large as possible to provide statistical power for future scientific discoveries.

Relative Value of Preconception, Prenatal, and Postnatal Data

A substantial literature, documented in the justification submitted by the NCS Program Office to the earlier review (National Research Council and Institute of Medicine, 2008), identifies prenatal infection, psychosocial factors, and environmental exposures as major contributors to child health that are in need of further study. The state of the art in longitudinal birth cohort studies is to begin data collection in early pregnancy: starting later risks bias from retrospective recall and an inability to measure transient, nonpersistent prenatal exposures from environmental or biological samples collected after birth. It also limits the ability to evaluate the role of the intrauterine environment as determined by various obstetrical conditions, such as hypertension, diabetes, or fetal growth restriction. The current NCS Program Office design is to enroll roughly half of its sample prenatally and half at birth.

Several major national birth cohort studies have enrolled women during the prenatal period and collected biological specimens at multiple points during pregnancy, including Generation R, a large population-based Dutch cohort (Jaddoe et al., 2012); the Norwegian Mother and Child Cohort Study, which

[5]Adjustment of sampling weights to reflect nonresponse and a host of more technical sampling issues will mean that after 21 years, sample members will not have equal weights. (See Chapter 3 for a more detailed discussion.) The most noteworthy example comes from the NCS's plan to include siblings in the probability sample, which roughly doubles the selection probabilities of second and later children born to the same woman during the birth window.

enrolled 108,500 children (Magnus et al., 2006); the Danish National Birth Cohort, which enrolled 100,000 families (Olsen et al., 2001); and the Japan Environment and Children's Study.[6] Multiple reviews of these studies have emphasized the value of prenatal specimen and data collection (e.g., Landrigan et al., 2006). The original NCS strategy of attempting to enroll families during the prenatal and possibly preconception periods was considered to be a strength by the earlier panel (National Research Council and Institute of Medicine, 2008).

Despite scientific consensus on the importance of beginning data collection during the prenatal period, the NCS Program Office did not provide a scientific rationale or a convincing financial argument to support the change in design to enroll fully one-half of the probability sample at birth. Furthermore, the panel found problematic the proposed approach of collecting environmental information through retrospective recall in interviews and assuming that samples collected during the immediate postnatal period can be used to characterize the pregnancy environment. (These issues are discussed in Chapter 4.)

The NCS Program Office justified its decision to split the cohort enrollment into prenatal and birth strata on the experiences of the Vanguard sites and the large incremental costs associated with the prenatal recruitment, enrollment, and data collection. Documents provided by the NCS Program Office stated that each prenatal enrollment and associated data collection would cost an additional $10,000 per recruited woman relative to a birth enrollment (NICHD, 2013b, p. 25). When asked to provide more information on this estimate, however, the NCS Program Office responded that the estimate was incorrect, but it did not provide a correction. Nor did the NCS Program Office provide a sufficient justification to the panel's request as to why this particular cost-saving strategy (splitting into prenatal and birth cohorts) was selected over other possible strategies. (Chapter 5 discusses the fielding cost implications of several alternative designs that would increase the size of the prenatal sample without significantly affecting the total budget.)

Considering the scientific framework and goal of treating the study as a platform for future research, the NCS should attempt to enroll as many participants as feasible during the prenatal period and to collect prenatal data as well as biological and environmental specimens from them. Enrollment at the time of birth should be limited to women who do not receive prenatal care or otherwise do not have a chance of selection through prenatal providers.

RECOMMENDATION 2-2: In order to facilitate scientific discovery during and after National Children's Study data are gathered, the Main Study should use a national probability sample with the largest feasible sample

[6]For a description, see http://www.env.go.jp/en/chemi/hs/jecs/ [May 2014].

size and an approximately equal probability of selection design, and it should recruit nearly all of the cohort as early in pregnancy as possible.

Quality, Standardization, and Flexibility in Data Collection Methods

A very large birth cohort study such as the NCS provides a unique opportunity to validate or confirm findings reported from smaller and more focused epidemiological studies, as well as to pool data with other large epidemiological studies to identify determinants of very rare conditions (such as specific childhood cancers). The NCS Program Office has engaged with the World Health Organization and investigators leading other national birth cohort studies in an effort to harmonize or coordinate the use of data collection methods to facilitate future pooling of data. To realize the potential of these opportunities, the NCS must use high-quality, well-validated, and standardized study methods and instruments. At the same time, the study will need to incorporate strategies to develop and validate new data collection methods to be able to address future domains of scientific inquiry. It may also be necessary to revise and shorten standardized instruments to reduce overall respondent burden while measuring many domains. Finally, the NCS needs sophisticated and well-developed computerized systems for data collection and management.

Quality control is key to the success of the NCS and adherence to protocols is critical. In the early Vanguard Study, NCS used a contractor to assure conformity in training of interviewers, instruments, and other aspects of the collection. NCS stated that there is also a plan to engage an independent quality control contractor as part of the Main Study (NICHD, 2013d, p. 55), and expanded briefly on that in NICHD (2013f, p. 5), saying that its experience is that "an independent assessment can improve quality, even with extensive quality control built into the process." In NICHD (2013d, p. 40), NCS also acknowledges the importance of high quality training for field staff. Cost effectiveness and efficiency are also key metrics of a high quality study. Little or no cost information relevant to the Vanguard Study was provided to the panel. However, the cost information for the multiple components of the Vanguard Study may not be directly relevant to evaluating future costs because the Vanguard Study spent significant time and resources to conduct large scale pilot testing of multiple recruitment strategies and data collection protocols. The Vanguard Study has or should yield relevant information for designing the Main Study, although some additional pilot testing may be needed to address gaps in information. The panel concluded generally that the NCS should not have to undertake complete full scale pilot testing of the recruitment strategies and data collection protocol on the scale of the prior recruitment pilots. Additional pilot testing that may be needed to address gaps should be designed to focus on obtaining the specific needed information using the most cost effective approaches.

In its commendable plans to coordinate data collected in the NCS and

other birth cohort studies, the NCS Program Office will benefit from guidance from its advisory committees, other governmental groups, and the scientific community on reaching the optimal balance between widely used standardized instruments and new or revised instruments. The NCS will also use the Vanguard Study to develop and validate data collection methods and instruments prior to their use in the Main Study.

An important strategy to maintain flexibility to address future scientific issues and to anticipate future innovations in measurement is for the NCS to collect and archive biological and environmental samples in ways that make them available for future investigations. The methods used to collect, process, and store samples should maximize the potential future use of analytical approaches (e.g., proteomics, metabolomics, genomics, transcriptomics), particularly when considering that some of these samples may be stored for decades. Archiving samples is also essential to reduce overall study costs, since future analyses can use nested case-control or case-cohort designs to limit the total number of samples that would have to be analyzed. Current NCS plans are compatible with this important design feature.

Essential in any data collection strategy is a robust system for tracking the information gathered, rapid coding of assessments to facilitate dissemination, and appropriate mechanisms for preserving and archiving biological and environmental samples for future use. To the extent feasible, the NCS should build on robust data management systems that have been developed for other large data collection efforts.

RECOMMENDATION 2-3: In order to facilitate scientific discovery during and after National Children's Study (NCS) data are gathered, the Main Study should use valid and standardized data collection measures and methods, while maintaining flexibility to revise or develop new instruments. The NCS should also use state-of-the-art procedures to collect, archive, and provide access to biological and environmental specimens for future analyses.

Dynamic Conceptual Framework

The earlier report (National Research Council and Institute of Medicine, 2008) expressed concern that there was no apparent overarching conceptual framework for health and development to tie the study together. In response, the current plan describes a detailed conceptualization of health and development. The breadth of the conceptualization would encompass most of the issues affecting child health and development and provide many dimensions that could be linked to environmental exposures, which should facilitate scientific discovery. This issue is discussed in more detail in Chapter 4.

A key component of the new conceptualization is that rather than measur-

ing specific diagnoses or syndromes, the Program Office plans to collect more detailed data on health status to allow future researchers more flexibility in defining health and disease phenotypes. The panel agrees that the flexibility to use data to generate a variety of phenotypes, rather than focusing on specific diagnoses, seems promising. However, as described in Chapter 4, the panel was not able to judge the overall merits of these new approaches because important details on the operationalization and effectiveness of these new approaches were not provided by the NCS Program Office.

RECOMMENDATION 2-4: The proposed strategy for the National Children's Study Main Study to collect detailed data on children's health status, conditions, symptoms, and behaviors should be followed to the extent possible, taking into account constraints of costs, operational feasibility, and the need to not overburden respondents.

SUPPLEMENTAL SAMPLES

Given the scientific value of the largest possible national probability sample, the panel carefully considered NCS's current plan to allocate a portion of its total sample to supplemental or "convenience" samples.[7] Specifically, the NCS Program Office proposes a probability sample of 90,000 births and supplemental samples of 10,000 births and seeks advice about the optimal composition of the 10,000. The plans as of October 25, 2013, were as follows (NICHD, 2003d, pp. 64-65):

> The only certain use of the [supplemental/convenience] sample is to enroll a cohort of preconception women enriched for those who are nulliparous to perform a preconception data collection visit and with the intent of scheduling a data collection visit as early in pregnancy as feasible during the first trimester. In the current proposal about 5,000 of a projected 10,000 births would be reserved for the preconception cohort. The use of the remaining 5,000 would not be defined until the Primary Sampling Units in the national probability sample are identified and characterized. Some part of the sample could be used for specific exposures that are of high scientific interest and public health value that were not included in the national probability sample. For example if none of the locations were located in an area that had fracking and there was sufficient interest and a scientific need based on a survey of other research efforts to collect data on possible exposures that occur near fracking sites, a location near a hospital with birthing services could be identified as a supplemental recruitment center.

[7]We prefer the term "supplemental" to "convenience" since some of the proposed samples could be drawn using probability sampling methods.

Other uses of the remaining 5,000 mentioned in the NCS briefing documents include subpopulations likely to experience disparities in health outcomes and not adequately represented in the 90,000-birth main sample; populations exposed to natural disasters, such as hurricanes or industrial accidents; and siblings born to mothers of enrolled children whose birth date occurs after the study's 4-year recruitment period. We consider each of these possible uses of the supplemental sample in turn.

Preconception Sample for First Births

Given the emerging scientific importance of prenatal and even preconception conditions for later health and development, the potential value of NCS information on preconception exposures could be quite high. The NCS Program Office proposes that the main probability sample include both its targeted births plus roughly 8,000 siblings of the targeted children born later during the 4-year birth window (see details in Chapter 3). Exposure information gathered before and after the birth of the target child provides preconception exposure data on the subsequent sibling birth. The panel strongly endorses this proposed sibling component of the main sample, in part because of the value of the preconception exposure data it will provide.

A preconception sample for first births is potentially valuable since preconception exposure information on first births cannot be gathered in the subsequent sibling portion of the main sample. According to the NCS Program Office, women at risk of becoming pregnant for the first time would likely be enrolled in the NCS supplemental sample through the recruitment of health care providers that offer health care services to nulliparous women. Working through these providers, the NCS would draw a convenience sample of the women most likely to become pregnant. These women would have an initial screening and home visit, during which environmental samples would be taken. Women would be followed at 3-month intervals by telephone:[8] if a woman becomes pregnant, she would be followed, using the prenatal and postnatal protocols in the Main Study. The Program Office believes that about 20,000 women would need to be recruited in this way to generate 5,000 first births.

In addition to the fact that preconception exposure information will be gathered for an estimated 8,000 siblings in the main sample who are born after initially enrolled subjects, the potential scientific value of the current plan for an additional preconception sample of first births would be not be high for two reasons: the proposed sample would likely not be representative of all first births, and it would incur high costs, including the costs of in-home interviews with four times as many women as are expected to eventually become

[8]A woman would be followed for as long as a pregnancy might lead to a birth during the 4-year birth window associated with the probability sample.

pregnant.[9] Given this mixture of benefits and drawbacks, the panel believes it is important to begin its analysis with an evaluation of the scientific case for including the first-birth preconception sample as part of the NCS sample.

Although fetal environments of first versus subsequent births may differ, the effects of preconception exposures to *persistent* environmental agents (i.e., exposures that persist between the preconception and prenatal periods) can be analyzed using the prenatal environmental information gathered for first births that occur in the main sample. The effects of important preconception exposures, whether persistent or not, that have similar effects on first and subsequent births can be analyzed using preconception data gathered from the sibling sample.

Consequently, a first-birth preconception sample provides uniquely valuable data only in the case of transitory preconception exposures that affect first births differently than subsequent births. None of the materials the Program Office provided to the panel referred to research showing such possible interactions. Thus, while preconception exposure information on first births may have potential to add scientific value to the NCS study, prior research provides no examples of such a value, and many of the possible links between preconception exposures and child outcomes can be investigated with the data to be gathered in the probability sample.

The panel also has a number of concerns about the design of the preconception sample and data collection. We note first that no details about provider and participant selection were provided to the panel, rendering a careful analysis of NCS plans impossible. For example, there was no mention of using probability sampling methods to select the nulliparous women. Even if the health care providers cannot be selected at random from the set of all providers, it would still be important to use probability sampling to select women receiving care from these providers. Given the need to develop and possibly pilot test[10] entirely separate recruitment and data collection protocols for the proposed preconception cohort and the hoped-for mid-2015 starting date for the study, the panel does not believe it is feasible to prepare the preconception sample for inclusion in the Main Study.

Second, the panel worries that insufficient steps would be taken to recruit nulliparous women who do not seek routine health care. Such women are important to include because they are most likely to be members of disadvantaged groups. Moreover, these women are most likely to be exposed to unhealthy environmental conditions of greatest concern for the NCS.

[9]See Appendix B for details on the likely field costs associated with a preconception sample.

[10]Pilot testing of recruitment of nonpregnant women through provider offices was done in the provider-based recruitment component of the Vanguard Study alternative recruitment pilot (see Chapter 1), but it was mostly limited to prenatal care provider offices and targeted nonpregnant women at high likelihood of becoming pregnant.

Third, since little is known about the types of factors or time periods prior to pregnancy that might be the most important, a study might need repeated and fairly extensive data collections to address this issue. Indeed, the original NCS strategy was to enroll nonpregnant women at high or moderate "risk of pregnancy" (National Research Council and Institute of Medicine, 2008) and collect preconception data multiple times for the women identified as high risk. The merits of the current NCS plan (one in-person data collection visit per woman, see NICHD, 2013d, p. 65) are not clear to the panel, and the NCS Program Office did not provide a scientific justification for this design decision.

The preconception data could be biased if, for example, environmental factors that potentially affect child health outcomes could also affect fecundity or time to pregnancy. Alternatively, fecundity could be an intermediate factor between an environmental factor and child health outcomes. There could be multiple complex causal pathways between the preconception environment and child health outcomes. It is doubtful that one preconception data collection of 5,000 women at varying times prior to pregnancy would be adequate to analyze these complex pathways.

Finally, the cost of the preconception sample of births is much higher than the cost of births in the probability sample. As detailed in Appendix B, the main reason for higher costs is that about 20,000 nulliparous women must be recruited and interviewed in their homes to yield 5,000 first births. Our cost analysis illustrates the opportunity costs of the preconception sample by showing that eliminating the preconception first birth sample would enable almost complete prenatal, rather than the currently planned half prenatal and half postnatal, recruitment of women and children in the main sample.

RECOMMENDATION 2-5: While the panel appreciates the potential scientific value of gathering preconception exposure information on 5,000 first-birth children as part of the National Children's Study Main Study, this supplemental sample should be dropped because of high costs, the lack of any evidence of the value of such a sample, the lack of detailed plans for both selection and analysis, and potential limitations in the proposed data collection schedule.

Supplemental Samples to Address Targeted Research Questions

The panel did not find any value of using the supplemental samples for the NCS's other stated purposes—namely, for populations living in geographic areas with possible exposures from conditions such as fracking, populations exposed to natural disasters such as hurricanes, younger siblings of enrolled children born outside the birth window, and augmented numbers of minority groups of interest for health disparities research. In part, this assessment reflects

the lack of detail from the NCS Program Office about the rationale for such supplemental groups and operational details.

In the case of the geographic exposure sample, the Program Office provided the panel with only the most general description of its plans and none of the details needed to evaluate them. This is problematic for many reasons. For example, the panel sees no reason why the locations for the geographic exposure sample cannot be identified in advance and included as strata in the Main Study probability sample. In addition, the needed coordination of sample design and study staffing between the probability sample and geographic exposure samples dictates that sample selection and recruitment for the geographic sample should begin at roughly the same time as sample selection and recruitment for the Main Study. Given the study's expected mid-2015 start date and that the Program Office has not yet identified the specific geographic exposures of interest that it would target, the panel fails to see how data collection for the geographic exposure sample could coincide with the data collection for the probability sample.

Moreover, the panel did not find any justification for devoting a portion of the NCS's sample to enrolling women in areas affected by meteorological, industrial, or other events of interest. In this case, the NCS would not be able to enroll women with births in the same geographic areas *prior* to the event, precluding scientifically strong pre- to post-event comparisons. Second, given the time it takes to set up sampling and interviewing mechanisms, the time between the event and actual data collection may be long. Third, for many events, a substantial proportion of affected women may have moved away from the affected area. And fourth, if post-event-only studies are to be conducted, the public availability of NCS instrumentation and other study protocols will make it possible for special studies to be mounted that focus more specifically on gauging the likely aftermath of the event on children.

A third possible use for convenience samples is to enroll younger siblings born outside of the 4-year birth window. We estimate that an expansion of the birth window from 4 to 7 years for these younger siblings only in each primary sampling unit would roughly double the number of siblings enrolled in the study (after accounting for an expected 20 percent attrition) to 18,000. However, given the plans to recruit about 8,000 siblings as part of the Main Study, the panel judges that the likely advantages of additional siblings do not outweigh the opportunity costs that expansion would entail, since resources needed to capture the additional 10,000 births could instead be used to accomplish other study goals, such as expanding the prenatal sample.

A final proposed use of the supplemental samples is to facilitate the investigation of subpopulations of interest for research on health disparities. The NCS's large probability sample should be sufficient to generate substantial numbers of children in the largest demographic groups commonly used in such research, as well as for subgroups (such as socioeconomic categories within

race and ethnic groups) necessary to properly investigate the extent of health disparities (see detailed discussion in Chapter 3).

RECOMMENDATION 2-6: The supplemental convenience samples proposed for the National Children's Study Main Study should be dropped from the design, including samples of children exposed to natural disasters or geographically defined environmental exposures, samples of additional members of disadvantaged groups, and samples of siblings born outside the 4-year birth window. The potential added value of the supplemental sample cases is less than the value of the additional cases in the probability sample they would replace, specifically, the value of the additional prenatal cases in the probability sample.

HEALTH DISPARITIES

Health disparities, defined as "systematic, plausibly avoidable health differences adversely affecting socially disadvantaged groups" (Braveman et al., 2011), exist for numerous health conditions and across people's lifespans. There are numerous examples of stark disparities for children, including: African American infant mortality rates are 2.5 times higher than white infant mortality rates (Hamilton et al., 2013); asthma prevalence for children is 2.4 times higher for Puerto Ricans, 1.6 times higher for African Americans, and 1.3 times higher for American Indian and Alaskan natives than for whites (Akinbami et al., 2009); and poor children are almost twice as likely as nonpoor children to have a serious health limitation (Seith and Isakson, 2011). The reduction and elimination of health disparities has been identified as an important goal by government agencies (e.g., in both *Healthy People 2010* and *Healthy People 2020*, produced by the Centers for Disease Control and Prevention), nonprofit groups, and community representatives.

Despite repeated documentation of child health disparities for many conditions, important questions remain about their fundamental causes. In response, the 2000 Children's Health Act directed that the NCS be designed to consider health disparities. The earlier report (National Research Council and Institute of Medicine, 2008, pp. 37-39) identified a number of deficiencies in the NCS's original approach to health disparities, including; (1) the decision to use equal probability sampling, which may lead to insufficient sample sizes for some racial, ethnic, and language minorities for some analyses; (2) low response rates in areas, such as inner cities, that are traditionally hard to survey and that will reduce effective sample sizes for disadvantaged groups relative to other groups; (3) lack of attention to generating data on how individuals from different groups may interact with health systems, a factor whose importance has been suggested in many previous studies; and (4) the absence of virtually any hypotheses about racial and ethnic disparities. The earlier study summarized

its concern as follows: "[w]hile the study will gather a great deal of information that is relevant to understanding such disparities, the research design was not informed by a concern with understanding their basis" (p. 5).

The NCS Program Office outlined several responses to these critiques (see NICHD, 2013b, 2013d). First, with regard to equal probability sampling, the Program Office noted that 90,000 of the 100,000 Main Study's cohort children would be drawn from a representative sample of hospitals and birthing centers, which collectively cover about 99 percent of U.S. births. By implication, subgroups of interest should be enrolled in the probability sample. Even in the absence of oversampling, the main NCS sample will contain thousands of children who belong to subgroups that constitute only a few percent of the overall U.S. population of births (see details in Chapter 3). Second, with regard to attrition, the Program Office noted (NICHD, 2013b, p. 10) that "The early Vanguard Study data indicate that a provider-based recruitment model demonstrates better response rates and retention rates than alternate models. In addition, NCS continues to invest resources in a comprehensive retention plan as called for in the 2008 IOM report." Third and fourth, in keeping with the NCS's belief that its scope "should be limited only by scientific creativity and not by current consensus priorities" (quoted in Guttmacher et al., 2013, p. 1873), the current study plan does not include any specific health disparity questions or hypotheses and does not address the concern about how different groups interact with health systems.

In summary, the NCS's approach to health disparities consists of four prongs: (1) ensure that populations of interest for health disparities research are adequately represented in the sample by including them in the probability sample and possibly using a portion of the planned 10,000 special sample for supplemental coverage of those populations; (2) ensure that information about the demographic and other characteristics that define these populations is gathered in the core NCS questionnaire; (3) ensure that exposures important for understanding health disparities are measured; and (4) devote resources to retain as many participants as possible in the Main Study.

Although the panel agrees that the large sample size and the comprehensive assessment of health determinants and health outcomes that is planned in the NCS will allow researchers to investigate many important health disparity questions, the relevance of health disparities to children and society, as well as the high importance of this topic to the NCS, requires that the NCS take special steps to ensure that the sample is adequate for addressing these questions. The panel's detailed analysis of these issues is provided in Chapter 3.

3

Sample Design

This chapter reviews the sampling-related elements of the proposed study design for the National Children's Study (NCS). After summarizing the proposed study design and its background, the chapter delineates established principles for evaluating the design of studies such as the NCS and then evaluates the proposed design against these principles. It ends by recommending next steps. Taken together, the chapter addresses statistical issues surrounding the following key items the panel was asked to consider: the national probability sample's overall sample size and design; the use of hospitals and birthing centers as the primary sampling unit; the use of health care providers to sample and recruit prospective participants; the relative size of the prenatal and birth strata in the probability sample; and the optimal use of sibling births.

PROPOSED SAMPLE DESIGN

As described in Chapter 1, the decision of the NCS to use probability sampling methods and the perceived importance of collecting preconception and early prenatal data led to an initial design of a national equal probability of selection sample of 100,000 births drawn from a stratified area sample design using geographic locations (mostly counties) as the primary sampling units (PSUs) and households within clusters of census blocks in the selected locations as the secondary sampling units (SSUs). However, the experience of the Vanguard Study indicated that a household-based sampling design was likely to be too inefficient regardless of the recruitment methods, so the NCS developed several alternative designs. Currently, the proposed design is a national equal

probability of selection sample[1] of 90,000 births using a stratified list sample design with hospitals and prenatal care providers as the sampling units and places of recruitment. The target population consists of all births in the United States (excluding U.S. territories) during a 4-year reference period. In practice, each PSU will possibly have different overlapping 4-year periods, the "birth window," during which births are enrolled in the study, because the start of the sample enrollment will be rolled out over time.[2] While the broad outline of this current sample design was provided to the panel, many crucial details of the sample design had not yet been resolved as this report was being written.[3]

The probability sample would use birthing hospitals and birthing centers as the PSUs and prenatal care providers whose patients deliver at the selected hospitals and birthing centers as the SSUs.[4] The list of providers associated with selected hospitals (the SSU frame) would be split into two strata: the details of this split were not provided to the panel. The first SSU stratum, the prenatal stratum, would consist of those prenatal care providers (practice locations) that would be sampled for inclusion in the prenatal stratum. This stratum would include births to women who were sampled and recruited at a selected provider at their first prenatal visit. The second SSU stratum (the birth stratum) would include women who had their first prenatal visit at a provider listed in the second provider stratum's frame. They would be eligible for sampling and recruitment into the study at the selected hospital shortly after delivery (NICHD, 2013d, pp. 13-14), as would women who had their first prenatal visit at a provider not listed in either stratum's frame (which could occur if a provider had too few patients to be listed, or a new provider location was established during the enrollment period) or who did not receive any prenatal care. Although providers are listed in the second SSU stratum's sampling frame, the providers in this stratum are not used as sampling units for the hospital-based sampling of births, though they are used to establish the eligibility of those births.

Although it would be possible to recruit all of the potential NCS participants who receive prenatal care during the women's initial prenatal visits, the

[1]The term "equal probability" is used somewhat loosely here because, as explained below, plans for the sibling component of the national probability sample will essentially double the selection probabilities of sampled siblings relative to other births. In addition, differential attrition and other, more technical, aspects of any sample design produce variation in the achieved selection probabilities in a planned equal probability sample.

[2]Children born to mothers who have previously agreed to participate in the study are enrolled at the time of their births. For women recruited into the study prenatally, the recruitment period will begin 6 to 9 months before the start of the birth window and will end 6 to 9 months before the end of the birth window.

[3]The proposed sample design also includes supplemental (convenience) samples, which are discussed in Chapter 2.

[4]The provider SSU is operationalized as practice locations, rather than practice associations. For example if ABC practice has two locations, and both locations deliver patients at the selected hospital, then ABC practice would be listed in the provider frame twice, once for each location.

Program Office stated that it would be too costly to enroll and collect data during the prenatal period for the full Main Study cohort. The current proposal is that 50 percent of births in the probability sample will come from the prenatal stratum and 50 percent from the birth stratum. The first- and second-stage sample design for the PSUs and SSUs, respectively, would be a stratified probability-proportional-to-size design using births in recent preceding years to develop the measure of size.

Because the PSUs will have more annual births than the target number needed for the NCS, individual women will be sampled. For the prenatal stratum, there will be sampling of eligible women at their "first prenatal visit" to a sampled provider. For the birth stratum, there will be sampling of eligible women who have a live birth at the selected hospitals and birthing centers. For both strata, the penultimate sampling stage will use randomly selected days and time periods at the sampled provider office, hospital, or birthing center. The ultimate sampling unit by which the NCS cohort is defined is technically the live birth, selected from pregnant women who will give birth or women who just gave birth during the birth window.

In the plan initially provided to the panel (NICHD, 2013b, 2013d), the sampling frame for hospitals was to be developed from the list of hospitals maintained by the American Hospital Association, augmented with a list of birthing centers from the American Association of Birthing Centers, and further augmented with other information such as natality data, the State Inpatient Databases, and other federal and commercial data bases. However, at the meeting of the panel in October 2013, the NCS proposed that the list frame be developed from the 2010 State Inpatient Databases of the health care cost and utilization project data repository compiled by the Agency for Healthcare Research and Quality of the U.S. Department of Health and Human Services.

Two months later, in December 2013, the panel received a summary from a preliminary report written by NCS expert sampling consultants that documented progress on design of the first-stage sample. Based on an initial analysis of data from 27 states and with no stratification, a sample size of between 200 and 300 PSUs (individual hospitals or hospital clusters) would "permit generation of reasonably precise national estimates of birth outcomes as well as allow for a nominal level of precision for analysis of relatively large subgroups (sex, larger race/ethnic groups, income quartiles)" (NICHD, 2013i, p. 2). The consultants' analysis used a size cutoff to omit hospitals with fewer than 50 births per year (birthing centers were not mentioned) in order to account for 99.9 percent of national in-hospital births.

The sample frames of providers would be constructed by working with each hospital selected in the first stage to identify a list of referring practice locations. Additional sources of information about prenatal providers would include: state licensing records, insurance lists, medical society listings, birth data from official state and county sources, and professional association mem-

bership lists (NICHD, 2013h, p. 1). The Program Office wrote that it may also administer provider questionnaires to use in preparing the sampling frame for providers (NICHD, 2013d, p. 71).

The NCS proposes to implement the study in phases by initiating the recruitment activities in different subgroups of the PSUs at different times. A 4-year sample rollout period,[5] when combined with the expected 4-year birth window within sampled PSUs, means that any given round of data collection (e.g., information gathered when a child is 6 months old) will last 7 years.[6]

Another novel aspect of the sample design (discussed in NICHD, 2013d) is to include in the sample with certainty all siblings born during the 4-year birth window to mothers with a child already in the sample (which we refer to as the "target" child or birth). This plan will allow for the collection of preconception and early pregnancy exposure biologic data for subsequent siblings that will not be available for the originally sampled (target) child, for whom data collection began later in pregnancy.

NCS also plans to monitor recruitment by category and increase enrollment efforts to achieve the desired representation (discussed in NICHD, 2013d, p. 80; illustrated with Vanguard Study data, pp. 6-8). The NCS also plans to follow movers (discussed in 2013e, pp. 4-5) and to target retention efforts at subpopulations at greater risk of attrition (discussed in NICHD 2013b, pp. 26-28).

EXPECTED ELEMENTS IN A SAMPLING PLAN

The panel evaluated the elements of the sampling plan described above against the kinds of sampling plans that are often developed for national surveys. Typical sampling plans describe a study's objectives and constraints and the steps proposed to operationalize and realize those objectives. Examples of surveys with readily-available sampling plans include the National Survey on Drug Use and Health,[7] the National Survey of Child and Adolescent Well-Being,[8] the Current Population Survey,[9] the National Longitudinal Study of Youth 1997,[10] and the National Postsecondary Student Aid Study.[11] These reports provide more detail than the NCS Program Office can be expected to

[5]NCS wrote that the "specific rollout plan is still under development" (NICHD, 2013j, p. 2). To be consistent with the cost model referenced elsewhere in the report, we assume a 4-year rollout period.

[6]Prenatal recruitment, of course, precedes the start of any birth window.

[7]For a description, see the appendixes at http://www.samhsa.gov/data/NSDUH/2012SummNatFindDetTables/NationalFindings/NSDUHresults2012.pdf [March 2014].

[8]For details, see http://www.ndacan.cornell.edu/datasets/pdfs_user_guides/092_Intro_to_NSCAW_Wave_1.pdf [April 2014].

[9]For a description, see http://www.census.gov/cps/methodology/techdocs.html [April 2014].

[10]For details, see https://www.nlsinfo.org/sites/nlsinfo.org/files/attachments/121221/TechnicalSamplingReport.pdf [April 2014].

[11]See the methodology report at http://nces.ed.gov/pubs2011/2011188a.pdf [March 2014].

BOX 3-1
Key Elements of an NCS Sampling Plan: An Illustration

A well-specified sampling plan for the National Children's Study would contain the following elements:

- Clear statement of the study objectives
- Clear statement of the target population that is to be sampled
- Target sample sizes at the beginning, middle, and end of the study
- For each stage of sampling in the prenatal and birth sample:
    - Definition of the population and sampling unit
    - Description of the sampling frame and its quality
    - Plans for stratification and allocation based on fixed and variable costs and variance
    - Target number of units to be sampled
    - Sampling protocol (equal or unequal selection probabilities within stratum)
- The sampling protocol for subsequent siblings
- A rigorous determination of the overall inclusion probabilities
- Expected completion rates at each stage
- Estimated design effects

It would also contain an explanation of how the sampling plan attempts to minimize both sampling and nonsampling errors.

specify at the design stage. Still, they suggest the type of information that is needed, and much of it can be specified before the first interview is conducted. All of those surveys' materials include a detailed description of the scientific justification for the various design decisions. For an illustration of how a comparable sampling plan would look for the NCS, see Box 3-1.

The NCS study objectives and main hypotheses need to be stated clearly and include important domains and key outcomes. The target population has to be fully described, and the sampling plan should delineate the target sample size at recruitment, birth, and key data collection milestones through age 21. For each sampling stage, the plan needs to provide a precise definition of the recruitment, sampling, data collection, and analytical units; and how and when the sampling will be implemented. The sampling protocol should be provided. In the case of the NCS design, the sampling protocol for subsequent siblings should be provided in detail. The plan should include a rigorous determination of the overall inclusion probabilities, expected yield rates, response rates, and retention rates at each stage (hospital, provider, birth) and cumulatively. Meth-

ods to adjust the target sample sizes for sample ineligibility and other sampling uncertainties should also be included in the plan.

The NCS sampling plan needs to be based on scientifically valid methods that attempt to minimize both sampling and nonsampling errors,[12] balancing the often competing goals of minimizing both variance and bias. Making the correct design decisions requires simultaneously considering many quality and cost factors, choosing the combination of design features and parameters that minimizes variances and biases while satisfying the specified costs and precision constraints.

EVALUATION OF THE NCS SAMPLING PLAN

The panel was asked to consider "the overall sample size and design" proposed for the NCS. The rest of this chapter provides our assessment of design elements when possible and indicates for what elements the panel has been unable to evaluate the design because the NCS Program Office did not provide enough information to do so. As described below and in Appendix A, the various NCS documents provided to the panel included needed details on some of these elements but not others. Relatively clear descriptions were provided of the target population, initial sample size, and plans for the inclusion of subsequent siblings. The incomplete nature of the hospital-based sampling plan generates many uncertainties about its quality and feasibility.

Target Population

The target population for the proposed design is all live births in the United States during a 4-year time period, referred to as the inference period in this document. Although they are part of the NCS target population, two types of births are excluded from the NCS sample frame: births to women who do not deliver in birthing hospitals or birthing centers and births to women at hospitals in which there are too few births to be included in the sampling frame.

The decision to use hospitals and birthing centers as the PSUs rather than geographic PSUs necessitates eliminating from the sampling frame the estimated 1 percent of births to women that do not occur in hospitals or birthing centers.[13] The interim proposed design (NICHD, 2013i) also provides a cost and operational justification for excluding hospitals with a very small number of annual births: namely, that it would be inefficient to establish and maintain

[12]This approach is referred to as minimizing total survey error (see Weisberg, 2005).

[13] The NICHD (2013e, p. 1) reports: "According to the 2010 and 2011 U.S. natality data, only 0.8% of women give birth outside of hospitals or birthing centers, mostly in homes. . . . These women are more likely to be US born, older, non-Hispanic white, married, and to have either less than 9 years or more than 16 years of schooling."

field operations for recruitment in such PSUs, and the interim design specifies a size threshold of 50 births annually for hospitals. Thus, any hospital with fewer births would not have a chance to be sampled. Planned future analyses will consider the feasibility of increasing the size threshold, which would increase the undercoverage.

All documents received by NCS before the interim proposed design document state that the sample is to be drawn from a list of hospitals and birthing centers. The design document does not state whether the same size threshold criterion would also be used for birthing centers, or even whether birthing centers are included in the database. If birthing centers are included on the same list frame with hospitals and a single size threshold is used, they are unlikely to be adequately represented in a probability proportional to size sample because they tend to have a much smaller number of births than birthing hospitals. If birthing centers are to be included in the sample, a separate list frame with a lower threshold might be considered; however, a separate sample of birthing centers may not be logistically feasible. Considering that birthing centers account for only 0.4 percent of births, with substantial variation by geography and race and ethnicity (MacDorman et al., 2014), it may be that the 0.4 percent of births at birthing centers should be excluded from the frame. Taken together and in light of these logistical considerations, the panel views these exclusions as reasonable, although the NCS plan for the inclusion or exclusion of birthing centers needs clarification.

Due to the possibility of seasonal patterns in births, exposures, and outcomes, the panel judges that the inference period should consist of full calendar years. Also, to maximize the utility of size measures that will be needed to sample with probability proportional to size at the PSU and SSU stages, the number of births to be used for those size measures should be based on whole years as well. However, the NCS plan is that the birth windows will be rolled out over time for the PSUs in the study. It appears that blocks of PSUs would begin enrollment in three or four approximately annual waves and there would be a phased activation of PSUs within waves over a period of several weeks to a few months, starting with the smaller PSUs first. Whether this rollout happens in blocks of PSUs being launched the first of each year or on a rolling basis over a few years, the different start dates for the birth windows will require reconciliation in the estimation process. Although this staggering of launch dates is necessary for a study of this size, it should be noted that these differences in start dates for the birth windows affect not only how the target population is described but may also increase the bias and variance of estimates because data from the birth windows will have to be adjusted to represent the inference period.

Deviations from a full 4-year birth window for any PSUs or sample components could also be problematic from an estimation standpoint. The NCS Program Office commented (NICHD, 2013i, p. 2) that "some of the larger

providers may have a slightly shorter than 4-year recruitment period." The panel's understanding is that larger providers would be fielded later due to their smaller sampling fractions and the flexibility that brings to the enrollment process, but that it would not follow that the duration of their birth window would be shorter. Similarly, the NCS Program Office's comment that enrolling women for the prenatal sample component would take place only through the first 3 years (NICHD, 2014a) is another possible example of such a planned deviation from a 4-year birth window. By definition, because the births must occur during a PSU's 4-year window to be eligible for the study, prenatal recruitment of births must start 6 to 9 months prior to the start of year 1 of the birth window and end 6 to 9 months prior to the end of year 4 of the birth window. We expect that women recruited during these sampling times who end up delivering earlier or later than expected, and who therefore do not deliver during the birth window, would later be considered ineligible for the NCS.

CONCLUSION 3-1: The panel endorses the proposed target population of all births in the United States during a specified time period consisting of 4 full calendar years, as well as the proposed sample exclusions from this target population.

Sample Size

An overall size of 100,000 first appears in NCS documents in 2002 and has been assumed or endorsed many times since then.[14] In the current proposal, the size of the probability sample has been reduced to 90,000. The NICHD (2013d, p. 20) states: [15]

> No matter what design NCS may use, there will be limits to detect associations between exposures and outcomes. The proposed design with a national probability sample size of 90,000 is limited to detect associations

[14]The second meeting of the NCS Advisory Committee in 2002 discussed a sample size of 100,000. For a record of that meeting, see http://www.nationalchildrensstudy.gov/about/organization/advisorycommittee/2002Jun/Pages/SAC_062002_minutes.aspx [March 2014]. In sample design documents presented to the NCS Advisory Committee in 2004, the sample size of 100,000 was a "given." See http://www.nationalchildrensstudy.gov/about/organization/advisorycommittee/2004Jun/Pages/other_work_062004.aspx [March 2014]. The design reviewed in the earlier study also specified a probability sample size of 100,000 (National Research Council and Institute of Medicine, 2008, p. 2).

[15]For the 30 hypotheses included in the 2007 design, NICHD (2013b, Appendix 2, pp. 40-41) illustrates minimum detectable odds ratios for various sample sizes (70,000, 80,000, and 90,000), for selected exposure percentages and outcome percentages under the previous (geographic-based) sample design. The table also includes the design effect ("DEFF") associated with the previous county-based sample design. A design effect is used to indicate the extent to which a sample design that deviates from simple random sampling increases the variance of estimates. In that table DEFF typically ranges between 3 and 4, but the maximum is about 20.

between exposures with a prevalence of about 3% and outcomes with a prevalence of about 2%.

In addition, a table in NICHD (2013g, p. 4) shows the sample sizes needed to detect effect sizes of various magnitudes with 80 percent power and 5 percent significance level[16] for different levels of exposure and different prevalence levels for the outcome. These calculations are based on an assumption of simple random sampling, although the document notes that, because a complex sample design will be needed, the sample sizes in the table will need to be "multiplied by design effects for the particular estimate under study." Because the hospital-based sample design has not yet been completed and actual selection probabilities for births have not been derived, these design effects are not yet known. The uncertainty surrounding the scope of the design effects makes it difficult for the panel to judge the adequacy of the proposed 90,000-birth sample size.

Another difficulty is the NCS Program Office strategy to deemphasize the use of explicit hypotheses to guide the study design so the NCS can serve as a platform for future research. This change makes it difficult to evaluate the proposed sample size by consideration of explicit study objectives. In particular, if lack of resources were to result in a smaller probability sample, analysis of specific objectives might become necessary to reevaluate the effects of the reduction on the study design, and, more broadly, to assess the feasibility of the study itself. The panel does endorse the notion of the NCS being a study platform, but this decision means the proposed design is likely to be inefficient and perhaps insufficient for certain types of analyses to answer specific questions that arise later.

CONCLUSION 3-2: Because of the lack of explicit hypotheses in the study design, it is not possible for the panel to judge whether the proposed sample size is justified on the basis of the study's objectives.

Equal Probability Sample

The stated goal of the NCS is to have an equal probability of selection sample to the extent possible.[17] The intent is to select each newborn, either

[16]In this case, the stated goal is to identify an odds ratio of 2 or greater with a power of 0.8 and a two-tailed significance level of 0.05.

[17]As stated by the NICHD (2013d, p. 11): A guiding principle of the NCS sample design is "simple sampling weights at the outset to increase utility of the data later." In addition, (NICHD 2013e, p. 3): "The NCS is proposing an equal probability design . . ."; and (p. 4) "the only circumstances when an unequal probability of selection may arise is in the sibling cohort." And in another document, the NICHD (2013g, p. 5): "We are currently planning on an equal probability sample of births, that is, not oversampling for any special group."

prenatally or at birth, with the same probability of being included in the sample, exclusive of the sibling sample. Despite the complexity and variations of the designs that have been proposed or investigated over the last decade, all have attempted to provide an equal probability of selection to all births occurring during the birth window. As discussed in Chapter 2, one rationale provided by the NCS is that "equal probability" sampling is a logical approach if the study is to serve as a study platform that would be able to address many current and possibly unanticipated domains of future scientific inquiry.

The previous review of the NCS (National Research Council and Institute of Medicine, 2008) judged that the lack of any oversampling of population subgroups is justified because the planned sample size of 100,000 would provide sample sizes for major demographic subgroups that are large enough to provide adequate statistical power across a number of subgroups of interest for research on health disparities.

The current panel reconsidered this issue, as well as whether oversampling might be needed to adjust for analytically interesting population subgroups, such as low-birth-weight babies to families of low socioeconomic status, who may have higher expected nonresponse and attrition rates. As noted above, the 2000 Children's Health Act directed that the NCS be designed to "consider health disparities among children." To address the issue of sample size for population subgroups important for research on health disparities, the panel calculated the fraction of births in 2011 for combinations of race, ethnicity, and maternal education level (factors related to low socioeconomic status). As shown in Table 3-1, one of the smallest percentages of births across the combination of these characteristics is 1.9 percent for non-Hispanic blacks with a bachelor's degree or higher. The NCS's initial sample size of 100,000 would be expected to yield about 1,900 such births without any oversampling.[18] The forecasted 20 percent attrition for the study would reduce this figure to about 1,500, which is still likely to be large enough to support many important estimates for this group.

Of course, the categories of "non-Hispanic Black" and "Hispanic" are broad and do not include other potentially important but smaller racial and ethnic subgroups, such as American Indians and Alaska Natives, Asians (overall and of different origins), or specific Hispanic national origins (e.g., Mexicans or Puerto Ricans). The sample design should take into consideration whether the benefits of including adequate representation of births in such smaller subgroups justifies stratification or oversampling, relative to the statistical cost of reducing the efficiency of the sample for making estimates of the overall population of births. More generally, the sample design needs to be explicit about the subgroups for which there may be adequate numbers and the subgroups for which the numbers may be inadequate. The lack of core hypotheses and

[18]Stratification could be used to help ensure this yield.

TABLE 3-1 Percentage of Births in Various Combinations of Race and Ethnicity and Maternal Educational Attainment, 2011

Race and Ethnicity	Less Than High School	High School Graduates	Some College	Bachelor's Degree or Higher	Percentage
Hispanic	12.1	7.6	3.3	2.0	25.2
Non-Hispanic White	6.0	15.1	13.0	20.0	54.1
Non-Hispanic Black	3.2	5.4	3.3	1.9	13.9
Non-Hispanic Other	0.8	1.7	1.4	2.9	6.8
Total Percent	22.0	30.0	21.0	27.0	100.0

NOTE: This table is based on the 80.0 percent of births for which both race and ethnicity and maternal school was known: N = 3,293,891.
SOURCE: Data from the 2011 Natality Public Use File, Centers for Disease Control and Prevention. See http://www.cdc.gov/nchs/data_access/vitalstatsonline.htm [May 2014].

research priorities to guide the design makes weighing such sample allocation decisions virtually impossible to resolve by standard scientific approaches.

Differential attrition is also an important consideration in deciding whether to use any oversampling in the design. While the stated intention of the NCS (NICHD, 2013b, pp. 26-28) is to monitor attrition by population subgroups and develop improved retention approaches when necessary, the panel is still concerned about disproportionate attrition among disadvantaged groups that have relatively more health and developmental problems. Most national longitudinal studies, such as the Early Childhood Longitudinal Study—Birth Cohort Study and the Fragile Families and Child Well-being Study,[19] have suffered disproportionate attrition among socially disadvantaged groups.

A strategy to counterbalance the analytic impacts of disproportionate attrition is to oversample disadvantaged groups at the beginning of a study relative to their likely attrition patterns through the middle or end of the follow-up period. This approach will increase statistical precision by both reaching and maintaining the targeted sample size for high attrition groups and by achieving final weights that are more homogeneous, thus reducing the effects of weight variation on the estimates.

During the panel's October 2013 public meeting, NCS staff expressed the belief that the advent of social media would likely produce different and less

[19]For details, see https://nces.ed.gov/ecls/birth.asp [March 2014] and http://www.fragilefamilies.princeton.edu/ [March 2014], respectively.

predictable attrition patterns in the NCS, making it unwise to alter the study design with disproportionate representation of disadvantaged subgroups. The panel believes that this judgment needs to be based on a careful review of the recent experience of U.S. studies with designs and survey contractors similar to those employed by the NCS.

CONCLUSION 3-3: By adopting an equal probability of selection design for the National Children's Study, it is likely that the sample sizes for a number of subgroups of interest will be inadequate for some important types of analysis. These subgroups are likely to include minorities in the U.S. population who are known to be on the negative side of health disparities and to have higher attrition in longitudinal studies. However, the absence of explicit study hypotheses and objectives makes it difficult to identify these important population subgroups and their associated sample size requirements.

Stratification

Stratification is a key element of most sample designs and has been mentioned but not detailed in documents provided to the panel. Stratification can help to ensure proportional representation of key subgroups of the target population and that the sample includes regions with varying levels of demographic characteristics, exposures, and other variables of analytic interest. It can also be used to oversample certain subgroups. In this section, we discuss the use of stratification to ensure variation of key attributes of subgroups to address the NCS's goals regarding health disparities and environmental influences.

Describing and understanding health disparities by race and ethnicity requires separating the independent contributions to health of socioeconomic status, race and ethnicity, and geographic location (see La Veist, 2005; Williams and Sternthal, 2010; Yang et al., 2004). Prior studies have had difficulty separating the effects of these different factors because of strong confounding of socioeconomic status and geography with race and ethnicity. The presence of sufficient variability in socioeconomic status and geography within race and ethnic groups is fundamental to answering key questions about health disparities.

The brief PSU design summary the panel received (NICHD, 2013i) indicated that the current PSU design is likely to be comprised of 200 to 300 hospital PSUs. That design, which did not include stratification, said that the contractors will continue to refine their analyses by adjusting some assumptions and evaluating potential stratification variables, including: area-level income, birthweight, infant death, race and ethnicity, premature birth, respiratory distress syndrome, and health insurance type; however, this work had not been completed as of February 2014.

To strengthen its ability to study health disparities, the NCS Main Study

sample could be stratified by characteristics currently thought to be strongly associated with these disparities, including the ones listed in the PSU design summary, as well as by geographic region, exposures, and urbanicity. The panel believes it is essential to stratify in such a way as to include in the sample a sufficient number of births into families with both high and low socioeconomic status within racial and ethnic minority groups, as well as assuring variation in exposures and geography.

Stratification can be done at more than one sampling stage, depending on what characteristics are available in the sampling frames at each level. In fact, the information needed to stratify may drive the NCS to one particular sampling plan over another. For example, if needed stratification variables are available at the county level but not for hospitals, that may be an argument for keeping counties as the PSUs, rather than using hospitals as the PSUs. Thus, those refining the NCS sample design should investigate the extent to which any or all of these proposed stratification variables (or variables highly correlated with them) are available for various sampling units, whether counties, hospitals, or provider locations.

The currently proposed plan to split the provider sample frame to enable half of the sample of births to enter the sample from prenatal providers and half from hospitals is an atypical form of stratification. While the panel recommends (Chapter 2) that the sample enrollment not be split into prenatal and birth strata, if this were to be implemented as planned, the Program Office would need to specify how the SSU provider frames would be split, including any stratification and probability methods, and the strategy for making the two subframes comparable.[20] If the split is nonrandom, then each half of the split would not be representative of the target population on its own and so an unbiased inference could proceed only when both subsamples are analyzed in combination.[21] This approach would be quite restrictive for some types of analysis—in particular, estimates using data collected prenatally.

RECOMMENDATION 3-1: The National Children's Study Main Study sample should be stratified by characteristics that will achieve variability in socioeconomic status within important population groups to support analysis of health disparities, as well as achieving variability in envi-

[20]The NICHD (2013d, p. 14) states: "In a balanced sample allocation model between birth and prenatal recruitment, the prenatal location strata frame will have a cumulative measure of the number of annual births needed equal to 50% of the PSU annual births."

[21]For example, if the birth stratum providers are not selected randomly, then any model that uses prenatal information as a predictor can only use half of the data, and that half is not a random sample. The degree of bias this will impart in an analysis depends on what nonrandom system was used for stratification. If, for example, mostly smaller practices were selected for the birth stratum and if richer women go to large practices, then the estimated relationships between prenatal factors and birth outcomes would be biased.

ronmental exposures and geography to support analysis of relationships between exposures and health outcomes.

Hospitals as Primary Sampling Units

The panel was asked to evaluate the proposed hospital-based sampling design, that is, the use of a list of hospitals and birthing centers as the PSUs. This approach differs from those tested in the Vanguard Study, which used geographically based PSUs.[22] It also differs from the approach proposed at the workshop on the design of the NCS Main Study in January 2013, which used geographically based PSUs with hospitals as SSUs (see National Research Council and Institute of Medicine, 2013).

In the locations involved in the Vanguard Study initial and alternative recruitment pilots, all hospitals at which women selected into the sample delivered were asked to participate in the study by providing birth specimens for women who had already consented to be in the sample. Hospitals were not used as sampling units. Hospital-based recruitment was tested in the Vanguard Study by targeting three hospitals selected in each of three geographic PSUs, asking them to collect birth specimens for all women in order to have the information for women recruited after birth. Currently, the Program Office has proposed using a list frame of hospitals as the PSU frame for the study. The stated rationale is that by selecting hospitals first, and then providers within hospitals, the study would minimize the number of hospitals to enlist to participate in the birth specimen collections.

The documentation provided by the NCS concerning the proposed sample design is incomplete because the statistical work to develop the first-stage sampling plan for hospitals had not been completed by February 2014. It may be that a list of hospitals could provide a useful frame for the NCS first-stage sample. However, the panel has very little information concerning the quality of the proposed frame as a basis for the proposed PSU design.

While the State Information Databases currently proposed as a frame appear to be a feasible approach, the panel is concerned about potential undercoverage because the analysis currently under way is based on data from only 27 states.[23] Because the objective of the NCS is to draw inferences that are representative of the U.S. population, all hospitals in the United States must be listed and have a non-zero chance of selection. If the state databases are incomplete, especially with regard to only some states, the Program Office would have to

[22]The previous set of 110 PSUs was selected by the National Center for Health Statistics in 2005, using the number of live births for 1999-2002 as the measure of size.

[23]According to the NICHD (2013i), data from 27 states are available without getting separate permissions from the states. The data from the remaining states may or may not be available to the NCS because the Program Office would have to obtain permission from each of the other 23 states to use their data.

consider a hybrid approach in which states that are fully covered in the State Information Databases hospital frame use the hospital-based sampling design and other states make use of a geographically provider-based design.

Furthermore, according to the documentation for the State Information Databases, the availability of variables varies by state, including the hospital identifier. There is also some need to verify that the distribution of births on which the sample will be drawn is consistent with expected births during the recruitment period. This factor is important because the health care environment is highly volatile, with patient populations shifting quickly among hospitals. If based on data from 2010, at least 4 years will have passed between the date of the sample and the start of the study. Hospitals could have joined different networks, changed affiliated provider groups, or experienced increases or decreases in the number of births of various types. Access to vital statistics data at the state level for the prior year would indicate if such shifts have occurred.

CONCLUSION 3-4: The panel has not been provided with sufficient detail on the planned hospital-based sample design and recruitment strategy to judge their merits and scientific validity or determine potential coverage bias and the availability of appropriate stratification variables.

The NCS Program Office has asserted that the proposed hospital-based sampling approach would make it easier to make substitutions for hospitals that refuse to participate or to replace those that are found to be ineligible for the NCS (for example, if they have closed the obstetrics service) relative to the county-based approach. However, no evidence was provided to the panel to support this claim.

In the Vanguard Study—as in many community-based studies—geographic sampling even at the county level, and especially at the segment level, could function as a method to achieve diversity in socioeconomic status, exposures, and race and ethnicity because they are clustered by geography. It is less clear whether and how a hospital sampling frame could identify a set of key variables that would allow a similar type of efficient stratification while still maintaining strict "all births" probability sampling.

Assessment of the proposed sample design, when completed, should include comparisons with the previous designs and variations to those designs. As noted above, a previous design, which was tested in three Vanguard Study locations, used geography to define the first-stage sample, with prenatal care providers selected from the sampled geographic areas, followed by recruitment of women from the providers. One of the challenges with this plan was the number of hospitals that would need to be enlisted to collect birth specimens for women already enrolled in the study. A variation, previously proposed by NCS (see National Research Council and Institute of Medicine, 2013), was to use the same sampled geographic locations, with the second stage to be a

sample of hospitals in the selected area, and the third stage to be the providers associated with selected hospitals. A comparison could also be made with a hybrid approach: using hospitals as PSUs in states for which a complete list is available and using geographic PSUs in other states. Any comparison needs to include a cost-effectiveness analysis of the options and an assessment of the ability of each option to ensure coverage and to control for such characteristics as race and ethnicity, socioeconomic status, age, and marital status so that the sample will support evaluation of health disparities.

If the NCS reconsiders a variation on the previous county-based PSU design, the PSUs would have to be redrawn to reflect more current data. However, maximizing the overlap between the old PSUs (at least the locations that had experience in the Vanguard Study) and newly drawn PSUs (see, e.g., Ernst, 1999) might add efficiencies because the birthing hospitals are already familiar with the NCS, and NCS contractors continue to follow the Vanguard Study cohort participants in these locations.

CONCLUSION 3-5: The panel has not been provided with sufficient justification for moving to hospital-based primary sampling units from the sampling approach previously proposed by the National Children's Study discussion at the 2013 NCS Workshop (see National Research Council and Institute of Medicine, 2013) and based on Vanguard Study pilot testing—namely, county-based primary sampling units with hospitals as secondary sampling units and providers as third stage sampling units.

Because the current plan calls for hospitals to be selected with probability proportional to size, it is important that a good measure of size be available for each hospital on the frame. Inaccurate size measures in a multistage sample can lead to a less efficient design. The best that can be achieved in practice is that such measures would be at least 1 year old at the time of sampling.

Better information concerning initial hospital cooperation rates and recruitment of women at hospitals after birth is needed for efficient protocol development, planning, and cost analysis. A very small sample of hospitals was used in the hospital sampling component of the Vanguard Study, and those samples were drawn within three geographic PSUs. The method of selection was not described to the panel, however, and in one county the "Study Center staff had an existing relationship with selected hospitals" (NICHD, 2013f, p. 2). A detailed plan for replacement of hospitals that decline to participate needs to be better delineated, as well as how such replacements will be dealt with in weighting and response rate calculations, especially if similar replacements do not exist or are otherwise not available.

NCS contractors will need to receive institutional review board (IRB)

approval from the NICHD IRB and possibly other IRBs[24] of record in order to receive individually identifiable information needed for the study from hospitals and providers. Based on Vanguard experience, most providers do not have separate IRBs. However, hospitals may have their own IRBs and require that the NCS use their IRBs for in-hospital data collection. This may lead to challenges in gaining cooperation, and may increase costs. (The panel's cost analysis did not include costs for IRBs.)

Providers as Secondary Sampling Units

The panel was also asked to evaluate the use of health care providers to sample and recruit prospective participants. The use of providers to recruit prospective participants was tested in the provider-based recruitment arm of the Vanguard Study (10 locations), and the use of providers to sample and recruit prospective participants was tested in the provider-based sampling part of the Vanguard Study pilot (3 locations). In both of them, providers were selected within the previous plan's geographic PSUs, and women were sampled within these providers. Even though many procedures that are being proposed for the Main Study have been tested in the Vanguard Study, the panel was not provided with detailed information about all approaches that were tested and how well they worked. Some detail and discussion of the NCS provider-based sampling experience has been published by former Vanguard Study principal investigators (see, e.g., Belanger et al., 2013).

Several aspects of the design are of particular concern to the panel: the lack of information concerning the feasibility of developing a complete sampling frame of providers (specifically, practice locations) for selected hospitals; definition of the appropriate measures of size for sampling; approaches for dealing with providers that use more than one hospital; and how women are proposed to be recruited at their "first prenatal" visit, including the precise definition of such a visit. Like the hospital frame, the provider frame would have to be complete, recent, and of high quality and contain, at a minimum, a size measure that can be used for a sampling probability proportional to size of providers. The plan for this stage of sampling is to stratify the provider frames "by such provider location characteristics as are available for that list. Stratification factors could include geographical location, size, race/ethnicity mix, percentage of women on Medicaid, depending on the data available" (NICHD, 2013h, p. 2).

[24]The NICHD (2014a, pp. 6-7) states "there will be no need for IRB approval at each hospital as the HHS Office for Human Research Protection (OHRP) has determined that the engagement of facilities in the NCS is not human subject research. Actual data collection will be performed by NCS contractors. NCS will make use of the central IRB at NICHD in accordance to the HHS Office of Human Research Protections. If a hospital opts not to conform to the OHRP determination and wishes to use the NICHD IRB, the NCS does not plan to pay the IRB administrative costs but will provide a standard submission package and annual reports."

As with hospitals, a detailed plan for replacement of providers at the time of recruitment needs to be better delineated, as well as how such replacements will be dealt with in weighting and response rate calculations. In response to expected attrition, the current plan is to select a replacement entity from the same stratum on the frame. After 2 years in the field, NCS plans to check the stability of the provider list and, if needed, draw replacement providers (from an updated frame) that are similar to those that withdrew from the study. The Program Office (NICHD, 2013g, p. 6) stated that, in the Vanguard Study experience, attrition of providers over the 1st year or 2 has been small: 1 of 49.

Births as Last-Stage Sampling Units

With their provider-based recruitment and provider-based sampling experiments, the NCS may have sufficient pilot testing experience with the sampling of women at prenatal provider offices. But little information was provided to the panel about the pilot testing of the sampling of women just after birth at hospitals. (Chapter 4 discusses concerns about recruitment and data collection of sampled women at hospitals.)

As currently proposed by NCS, some births will be sampled at hospitals whether the hospital sampling is used only for women who have had no prenatal care or who received their first prenatal visit from a provider not included in the provider frame (as recommended in Chapter 2). While some of the procedures for identifying women eligible for hospital recruitment have been tested in the Vanguard Study, the panel was not provided with information on how these hospitals were selected, only that some may have been positively disposed to the NCS investigators.[25]

For hospital-based sampling, it is important that the NCS demonstrate it is able to: (1) identify all women eligible for selection in that setting; (2) sample them with known probabilities that can be adequately controlled during the recruitment process; and (3) recruit them with high success rates before these women leave the hospital. It is not clear to the panel that all hospitals will be willing to allow women to be recruited shortly after delivery or to assist in the identification of women in labor who are eligible for sampling in a consistent and scientifically valid way. It is likely that the opportunity to sample and recruit some eligible women will be missed. To calculate selection probabilities and response rates, it is crucial that all deliveries that are eligible or potentially eligible for inclusion in the study be accounted for, even if the recruitment opportunity is missed.

The Vanguard Study conducted a limited hospital-based sampling approach (proposed for the birth stratum). Attempts were made to recruit three hospitals

[25]In one location, researchers had existing relationships with hospitals and were able to make use of the hospital's electronic records to prescreen women for eligibility (NICHD, 2013f).

in each of three locations, and seven of the nine participated: thus, the provider-based sampling recruitment rate for hospitals was 77.8 percent.[26] However, the Program Office pointed out that the two that did not participate were not refusals but that the project ran out of time so it is not possible to predict what the ultimate refusal rate would have been. Based on information from all hospitals approached during the Vanguard Study, the Program Office has stated that they "they expect no more than 15%, and perhaps no more than 10% of hospitals to decline to participate" (NICHD, 2013d, p. 56).[27] However, this may be optimistic because most of the hospitals in the Vanguard Study were asked only to collect specimens for women who had already consented to participate in the study—a much lower level of effort than what is currently proposed for the Main Study. If the study design includes a substantial hospital-based sampling stratum, a pilot study would be needed to evaluate these issues.

The Possibility of Nonresponse Bias

Based on the Vanguard Study, nonparticipation rates at each stage of sampling (hospital, provider, pregnant women, and births) and the associated cumulative nonparticipation rates appear to be high.[28] Aside from nonparticipation of hospitals and providers,[29] which will be addressed using sample substitution and perhaps not included in the denominator of the response rate, it appears that provider-based sampling of women in the Vanguard Study (NICHD, 2013b, p. 14) had the following response rates:

- contacted for screening among sampled: 74 percent
- completed screening among contacted: 70 percent
- recruited among screened and eligible: 68 percent[30]

[26]The panel calculated this rate from NICHD (2013f, p. 4): 7/9 = 77.8 percent. This is rounded in NICHD (2013d, p. 72) which states: "In PBS [provider-based sampling] experience the hospital recruitment rate was 80%." The panel's cost analysis assumed that 20 percent of hospitals would agree to participate.

[27]There is also considerable local variation. In one Vanguard Study location that implemented provider-based recruitment, Kerver et al. (2013) found that the participation rate among hospitals was 71 percent, even though hospitals were only engaged to obtain biospecimens for previously consented women.

[28]The NICHD (2013g p. 7) agreed with this statement but clarified that "the response/participation rates at each stage of the Vanguard Study have been acceptable and comparable to or higher than other surveys of this nature."

[29]Results from the provided-based recruitment and provider-based sampling in the Vanguard Study indicate expected provider recruitment success rates of 64 to 68 percent.

[30]The product of rates in the above bullets for "completed screening among contacted" and "recruited among screened and eligible" is 47.5 percent, an estimate for the percentage of women contacted who were eligible and recruited. The panel's cost analysis assumed that this rate is 50 percent.

The response rate at recruitment (i.e., the fraction of women recruited relative to the total number of women sampled and estimated to be eligible) is 35 percent—the product of these three response rates.

This low response rate is problematic in part because if it differs across groups, the resulting estimates based on the NCS Main Study sample may be biased, even with well-designed weighting adjustments. Findings from the initial Vanguard Study recruitment pilot showed variation in cooperation rates at each stage across PSUs and by race and ethnicity, with significantly lower consent for women eligible for recruitment among Asians (Baker et al., 2014). The current design also may introduce differential enrollment between minors and adults due to the requirement for consent of the legal guardian for unemancipated minor pregnant women, because the legal definition of an emancipated minor varies among jurisdictions.

Some of the children enrolled in study will be lost to attrition, which is likely to occur differentially across groups.[31] According to NICHD (2013b, p. 27): "[The] NCS Vanguard Study data shows for preconception and prenatal women attrition between enrollment and birth ranges from 10 to 20 percent."[32] The Vanguard Study experiences were similar to those reported by other studies: attrition is greater during the initial one or two data collection visits after enrollment and then tends to decrease. In the proposed birth stratum, if this pattern occurs, there likely will be more substantial attrition during the first six months after birth than will be observed during the six months after birth in the prenatal stratum, as the latter already experienced their initial attrition pattern during their prenatal period. However, data to estimate attrition during the first six months after birth in the birth stratum are not yet available from the Vanguard Study (see NICHD, 2013f, p. 2). Attrition due to mobility[33] is greater in urban areas and among populations of low socioeconomic status who live in rental housing. In addition, NCS estimates that the annual retention after birth will be between 97 and 99 percent based on the experience of other longitudinal studies (NICHD, 2013b, p. 26).[34]

It is difficult to estimate the cumulative fractions of women and children likely to be lost over the course of the study. The assumed 10 percent to 20 per-

[31]By attrition, we mean individuals who drop out of the study after recruitment. In a study as complex and long-running as the NCS, there are also challenges because sample members may provide incomplete information on survey or other data collection instruments. These kinds of missing data issues are not addressed here.

[32]Some of this sample loss is due to ineligibility (e.g., miscarriage, moving out of the PSU) rather than attrition.

[33]Though the NCS plans to follow movers post-birth, many of them are likely to be difficult to follow.

[34]The panel's cost analysis assumed 10 percent attrition in the first year after recruitment of the mother, 3 percent in the second year, 2 percent in the third year, and then 1 percent annually after that.

cent attrition rates between prenatal recruitment and birth reduce the 35 percent response rate at recruitment to a cumulative response rate at birth of between 28 percent and 32 percent.[35] Even an annual 98 percent retention rate between birth and age 12 reduces this estimated range of cumulative response rates to between 22 percent and 25 percent. Maintaining a 98 percent annual response rate through age 21 produces an estimated cumulative response rate of between 18 percent and 21 percent.

CONCLUSION 3-6: Assuming that participation in the National Children's Study Main Study follows patterns in the Vanguard Study, the cumulative response rate to birth for the prenatal stratum would be between 28 and 32 percent, and the rate to age 12 would be 22 to 25 percent—very small fractions of the eligible sample. The cumulative response rate to age 21 would be 18 to 21 percent. A thorough analysis of nonresponse bias is clearly indicated, and in any case will be required by the U.S. Office of Management and Budget.

CONCLUSION 3-7: The high rates of cumulative nonresponse expected in the National Children's Study pose a severe risk for nonresponse bias that may not be mitigated by weighting adjustments, potentially making some study results invalid.

Optimal Use of Sibling Births

The current plan includes enrollment of all siblings from multiple births and all siblings born subsequently to the originally sampled "target" children but within the 4-year birth window. Assuming 20 percent cumulative attrition, there should be approximately 2,000 of the former and 8,000 of the latter.[36] The panel endorses this proposal (as discussed in Chapter 2), but we note some issues with the plan.

Enrolling subsequent siblings in this way offers some important analytical advantages, but it also entails costs and loss of precision. Foremost among the advantages is that the sibling data will provide information on the preconception and very early prenatal maternal environments of mothers of the study's target children who have a subsequent birth during the enrollment period. Such measures are not available for the target births since women are recruited into the study, at the earliest, at their first prenatal visit.

[35]Because the 10 to 20 percent attrition between prenatal recruitment and birth includes loss due to both nonparticipation and ineligibility, this cumulative rate is technically no longer a "response rate" as defined by industry standards. We use the term "cumulative response rate" here to indicate the rate of continued participation in the study among the population assumed to be eligible at the time of prenatal recruitment.

[36]These figures were estimated independently by the panel and the NCS Program Office.

Another analytic advantage of collecting data on siblings (including multiple births) is that sibling data can be used to estimate so-called sibling fixed-effect models that relate sibling differences in outcomes of interest to sibling differences in early environments (see Wooldridge, 2012). These models will allow researchers to control for characteristics of early environments that siblings share and that do not change over time (e.g., elements of maternal background), reducing bias in the estimated effects of variables of interest.

One analytic disadvantage of including siblings in the sample is that siblings are more alike than children chosen at random from the population, which reduces the precision of statistical estimates.[37] Another drawback is that the preconception and early pregnancy data collected on these subsequent siblings cannot be generalized to all births because, while originally sampled births are a mixture of first and higher-order births, none of the subsequent siblings is a first birth.

CONCLUSION 3-8: Enrolling siblings as members of the National Children's Study sample provides many analytic advantages, most prominently the gathering of preconception exposure information for second- and higher-order births. The panel endorses current plans to recruit siblings born after the initially recruited child—but only within the 4-year recruitment interval associated with the original primary sampling unit for the target birth—and to continue to follow these children until age 21.

Because subsequent siblings are to be included in the NCS with certainty, their probability of selection is quantifiable and they can, therefore, be considered part of the probability sample. For a probability sample of 90,000 births, this means that about 80,000 would be comprised of target births (recruited at providers or hospitals). About 2,000 would be multiple births (e.g., twins) and 8,000 would be subsequent sibling births.

The NCS has proposed at least two options for dealing with the fact that the subsequent siblings have more than one chance of selection in the sample: using multiplicity weighting adjustments to the sampling weights and screening a woman at the time of recruitment to determine whether she had a prior birth during the enrollment period for that PSU. In the second option, if a screened mother indicates a prior birth during this period (using one of the providers or hospitals listed on the frame), the child associated with this mother's current pregnancy already had a chance of selection as a subsequent sibling and would,

[37]For example, based on data from the National Longitudinal Study of Adolescent Health, Duncan et al. (2001) found that the correlation between same-sex full siblings in adolescent receptive vocabulary is about 0.50; in height it is about 0.46; and in an index of delinquent behavior it is 0.25.

therefore, be screened out of the sample, because its older sibling had a chance of selection into the study as a target child (whether actually selected or not).[38]

The weighting adjustments (option 1) could affect the precision of estimates for the total sample (90,000), in the same way that oversampling demographic groups could introduce weight variation, making them less precise. The consequences of these multiplicity adjustments are unknown but could lead to reductions in statistical precision and power. The panel has not been provided information that the screening methodology (option 2) has been pilot tested.

In either case, detailed information about all births to a woman within the enrollment period (date and place of birth, date and place of prenatal visit) has to be routinely collected at the time of sampling and recruitment. Option 1 requires additional effort to compute selection probabilities for prior births, and option 2 could substantially increase the recruitment effort and time needed to obtain the targeted number of women in the Main Study sample. Pilot testing could be used to develop procedures and questionnaires for both options, determine how well they work, and estimate how many women would be screened out under option 2.

CONCLUSION 3-9: Weight adjustment and screening are viable options for accounting for the fact that subsequent siblings have more than one way to enter the sample. The panel was not provided sufficient information to recommend one over the other. In either case, detailed information on prior births to the mother will need to be collected.

In terms of analytical issues, the NCS needs to consider how the subsequent siblings will be combined with the target children to make national estimates. One possibility is for the estimates to be stratified by birth order. Design-based estimates of preconception and early pregnancy exposures can be made from the subsequent siblings alone, but as mentioned above, these preconception and early pregnancy findings cannot be generalized to all births, only to second and higher-order births.

NEXT STEPS FOR THE SAMPLE DESIGN

Overall, the documents given to the panel did not provide sufficient details for an evaluation of whether the proposed sample would meet the minimal standards of a carefully specified, scientifically based sample design required for large national data collections. Many of the NCS design changes since 2008 appear to have been reactive and piecemeal, in response to issues that have arisen during the field testing or from prior reviewers, rather than from

[38]The likelihood of more than one child from the same mother actually being selected as part of normal sampling procedures at selected providers or hospitals would be quite small.

a planned and comprehensive approach to the design. One key missing document is a hospital-based PSU design with an assessment of the quality of the frame, discussion of stratification, selection of PSUs necessary to achieve design targets, methodology for replacing hospitals that decline to participate, details concerning calculation of selection probabilities, and proposed weight adjustments for nonresponse and attrition. While the NCS does have substantial pilot experience with the county-based PSU design in the Vanguard Study, little equivalent experience exists for the hospital-based PSU design.

CONCLUSION 3-10: As of February 2014, the currently proposed hospital-based sample design for the National Children's Study had not been sufficiently developed or documented to support an evaluation.

CONCLUSION 3-11: The identification, sampling, and recruitment of women at the time of birth have not been sufficiently pilot tested, using a representative set of hospitals, to support any conclusion about this feature of the design of the National Children's Study.

A detailed sampling plan and recruitment strategy for the NCS needs to be fully developed and documented by sampling and survey experts who have extensive experience in conducting large longitudinal national surveys. This group should be external to the Program Office but would work in collaboration with it on all aspects of the design. The sampling plan needs to include a justification in greater detail for moving away from the geographically based provider-based sampling tested in the Vanguard Study to the currently proposed hospital-based design. The plan also needs to address the need for oversampling various subgroups that are expected to have a higher attrition rate over the life of the study. The stratification plan may, in turn, dictate whether a geographic- or hospital-based PSU is used for the NCS, as stratification variables may only be available for one of these two PSU types.

A plan to mitigate nonresponse bias also needs to be developed before the study moves forward, and it needs to include identifying and collecting auxiliary variables or covariates that are thought to predict response and retention propensity. Outside survey experts can help determine: what metrics to use to monitor the extent of bias in survey results; what reporting strategies to use to monitor those metrics during data collection; and how the covariates will be used to adjust analysis weights to mitigate nonresponse bias.

RECOMMENDATION 3-2: A detailed plan for sampling, recruitment, and minimizing attrition bias for the National Children's Study's (NCS) Main Study should be fully developed and evaluated by sampling and survey experts independent from the NCS and approved by the proposed independent oversight committee before the study moves forward.

The NCS needs to evaluate and document what has been learned from the Vanguard Study. The independent sampling and survey experts, in collaboration with NCS, should determine what further pilot testing may be required. The NCS will need to conduct any needed pilot tests[39] or otherwise demonstrate the ability to carry out each stage of sampling, recruitment, and initial data collection. The sample in the pilot test needs to be large enough and essentially representative so that estimates of coverage, cooperation, and other key rates that affect costs and sampling validity can be made. To the extent that some aspects of this design have already been pilot tested, this information should be analyzed to identify gaps that require pilot testing. Results from the Vanguard Study and other pilot testing can be used by the independent survey experts in determining the final study design. We also recommend an oversight committee (see Chapter 6) that would approve the final design before the Main Study is implemented.

[39]The NCS Program Office has indicated that it is committed to full pilot testing of the alternatives considered for the Main Study.

4

Study Content

This chapter reviews the conceptual framework for designing the study content and evaluates the proposed study visit schedule, the proposed collections of data and samples, and the approach to define and characterize health disparities. It addresses two of the key items in the panel's charge: the proposed study visit schedule with its emphasis on more frequent data collection in pregnancy and early childhood and the proposed approach to define and characterize health disparities. The chapter also comments on the approach to data dissemination after the study is under way because an important objective of the National Children's Study (NCS) is to make the data publicly available.

The panel did not receive specific study protocols, information on specific data collection methods, or study instruments. Consequently, this review cannot address the scientific merit or quality of these aspects of the NCS data collection.

STUDY CONTENT CONCEPTS AND PARAMETERS

As noted in Chapter 1, several NCS features in addition to the sampling frame have changed since the previous review (National Research Council and Institute of Medicine, 2008). The current proposal conceptualizes the study less as a vehicle for testing current hypotheses and more as a platform for future researchers to formulate and investigate hypotheses that could be tested using data from the study instruments and previously collected biological and environmental specimens. Consequently, the proposal no longer relies on specific

pre-specified hypotheses to define the content of the study. Instead, "the proposed plan was developed using several exemplar hypothesis so it is hypothesis informed but not hypothesis limited" (NICHD, 2013b, p. 34).

The NCS is also now conceptualized as a platform for researchers covering a broad range of health domains. It will not focus on classifying participants into predetermined disease categories, but instead will collect a set of primary observations and events to enable researchers to apply their own health criteria and form "cases" (p. 31). According to the NCS Program Office, this framework can be used to develop more flexible phenotypes. A theoretical framework of health as multidimensional and dynamic will guide the selection of assessment methods and instruments.

Study Platform and Exemplar Hypotheses

The panel recognizes that not all cohort studies, particularly of this size, are designed around specific hypotheses. Some, like the Norwegian Mother and Child Cohort Study, appear to use the "platform" approach of the NCS or follow specific precedents, such as the Millennium Cohort Study and earlier British birth cohort studies. However, more commonly, studies are organized around key questions or assessments of specific exposures and provide details on specific subquestions and how they inform the data collection. While there are a variety of approaches to delineating the design and content of a study, most studies appear more focused than the NCS, for example:

- The Generation R study focuses on five specific areas: see Hofman (2004). Each area has one or more "aims" with more specific questions linked to that aim. Those questions drive the data collection for that aim.
- The Fragile Families and Child Wellbeing Study relies on four overarching questions.[1]
- The Norwegian Mother and Child Cohort Study seems to take the platform approach of the NCS: see Magnus et al. (2006).
- The Millennium Cohort Study follows the precedent set by prior British birth cohort studies; specific questions are not listed.[2]
- The French Elfe Child Cohort study has a list of seven research questions to be answered by the study: see Charles et al. (2011).

[1]For a description, see http://www.fragilefamilies.princeton.edu/core.asp [April 2014].

[2]For a description, see http://www.cls.ioe.ac.uk/page.aspx?&sitesectionid=880&sitesectiontitle=Survey+Design [April 2014].

A summary of several of these studies and others indicates that the studies usually address broad, but focused questions that are not termed "hypotheses."[3]

While the strategy to develop exemplar hypotheses or to state domains of interest rather than specific aims seems to be consistent with the range of approaches used in other large birth cohort studies, the panel determined that it was important to evaluate the specific strategy to use exemplar hypotheses proposed by the NCS. The panel reviewed material provided by the Program Office and asked for additional information about the exemplar hypotheses in order to understand how the hypotheses will be used to guide the study design. NICHD (2013b) mentions study hypotheses only twice, stating that the "proposed plan is hypothesis informed because it was developed using modeling of several exemplar hypotheses, but the plan is not hypothesis limited" (p. 5) and that the proposed plan was developed using several exemplar hypotheses (p. 34). The document included an appendix related to sample size that listed the hypotheses from the 2007 research plan, but it did not list the current exemplar hypotheses.

At the request of the panel, the Program Office later provided a description of the exemplar hypotheses. The document (NICHD, 2013d, p. 45) stated that the development of hypotheses "included a matrix approach utilizing exposures at various prevalence levels and outcomes at various prevalence levels as well as individual exemplar hypotheses." The document listed five examples of exemplar hypotheses based on four "exemplar exposures" and four "exemplar outcomes." None of the exemplar hypotheses specified time periods for the relevant exposures (e.g., first trimester, puberty), although one mentioned cord blood as a biological sample. None of the hypotheses mentioned assessment of confounding, effect modification, or gene-environment interactions. Other than mentioning that hypotheses were used as a general guide to estimating sample size, the NCS documents did not clarify how the exemplar hypotheses informed key decisions regarding the study design.

The sample size discussed in the document focused only on main effects, although effect modification or interactions is an important justification for a sample size in the range proposed for the NCS. Nor do the hypotheses quantitatively address implications of study power for assessing transient and nonpersistent risk factors (e.g., transient chemical, social, or maternal uterine conditions). In general, the sample size calculations were based on outcomes with prevalences at 2 percent and exposures at 3 percent: it is not clear how less prevalent conditions (e.g., autism, malformations) would be analyzed even though they are included in the document (pp. 44-45). The Program Office's response to a query on this point (p. 44) was a general disclaimer that the NCS could not address all important children's health issues.

An important criterion in developing exemplar hypotheses should be

[3]For a description, see http://www.birthcohorts.net/bch2/?action=list [April 2014].

whether the hypotheses anticipate the possible trajectory of future scientific inquiry in children's health and development. At the same time, it is also important to recognize that a unique strength of the NCS should be the capability to resolve inconsistent findings reported by smaller or more focused birth cohort studies. There have been a large number of smaller birth cohort studies, but the findings have not always been consistent, and important scientific issues are not yet resolved because of the limitations of the smaller studies. Addressing important unresolved scientific issues identified by recent epidemiologic research on children's health and development would be an important criterion to be considered in developing exemplar hypotheses for the NCS.

While exemplar hypotheses can serve as meaningful archetypes of important scientific issues that the NCS should be able to address, the few exemplar hypotheses provided to the panel are not sufficient to serve as the primary basis for planning a long-term birth cohort study. The Program Office described additional planning strategies in materials provided to the NCS Federal Advisory Committee (NICHD, 2014b, p. 7), including the concept of anticipating the developmental trajectory of a healthy 21-year-old person and then identifying "potential antecedent factors that could be measured earlier in assessing exposure that contribute to later outcomes." Another strategy is to identify multiple use cases, which would be scenarios of "sequences and interactions related to a particular outcome. Multiple use cases can be used to frame specifications for a system such as a research study" (p. 3). However, information on these additional strategies was not given to the panel, so the panel cannot evaluate the overall effectiveness of the study planning. Even if multiple study planning tools will be used, the NCS plan still must delineate specific scientifically robust exemplar hypotheses that can be used by the NCS and the scientific community to formally evaluate sample size and design issues, as well as the NCS proposals for the nature and timing of data collection.

CONCLUSION 4-1: A strategy of using a few exemplar hypotheses rather than stating a large list of hypotheses requires that the planners of the National Children's Study ensure that the exemplar hypotheses are important and scientifically robust to guide the study design and data collection.

RECOMMENDATION 4-1: Prior to proceeding with the Main Study, the National Children's Study (NCS) should develop scientifically well-grounded exemplar hypotheses that should be used to guide and evaluate decisions regarding the NCS design and data collection schedule and domains.

RECOMMENDATION 4-2: Because hypotheses will change over time, the National Children's Study should implement a strong and public pro-

cess to revise and develop new exemplar hypotheses to guide future study implementation, engaging with the extramural and intramural research communities.

Health Phenotype Concept

The NCS proposes to use a health phenotype and profile to describe each participant (NICHD, 2013b, p. 30):

> The term phenotype is used for the observable characteristics including morphology, physiology, developmental stage, behavior and products of behavior. . . . The term, profile, is used for the larger concept of phenotype plus environmental context. A profile includes observable characteristics about the participant plus information about the environment such as air particle measures, noise level, family structure and dynamics, access to health care, etc.

Thus, at each study visit a participant will be assessed using a health phenotype framework and will be the subject of collection of environmental data and biospecimens. Documents provided to the panel state that the rationale for using a health phenotype concept is that it would (NICHD, 2013b, pp. 29-30)

- Use a conceptual framework grounded in health that applies to all Study participants.
- Capture a broad scope of outcomes and not limit observations to particular conditions or diseases.
- Establish consistency in reporting exposures and outcomes across different research fields that may have different paradigms and methods.

The document also states that using the health phenotype concept would achieve another objective (NICHD, 2013a, pp. 21, 23):

> . . . maintain flexibility as new opportunities and assessment innovations arise as they can be integrated into the conceptual framework. . . . Using reactive airway disease as an example, the NCS will emphasize accurately capturing medical history, participant experiences, and respiratory symptoms, coupled with biospecimens, genetic analyses, and environmental samples. Researchers can then use these data in conjunction with the case definitions and classifications they deem optimal for their analyses.

In the information provided to the panel, the health phenotype concept and the conceptual framework for health development (which are discussed below) are said to guide "the development of assessments and the structure of data collection to ensure that essential relevant information to understand

health and development are included" (NICHD, 2013b, p. 31). The NCS documents provide examples of possible study visit content, but they do not explain how the health phenotype concept and conceptual framework will guide specific decisions about the content (see, e.g., NICHD, 2013b, App. 3).

Although the health phenotype concept is consistent with the overall NCS strategy to serve as a platform for future research, the panel cannot evaluate the scientific merit or ability of the NCS to operationalize the concept on the basis of the limited information provided to it. A specific concern is that the documents do not adequately explain the criteria and procedures that will be used to prioritize data collection, considering that the volume of detailed data collection needed to implement the concept could result in substantial respondent burden.

Consider the example provided by the Program Office for reactive airways disease: it could require a very large number of questions, including about symptoms, activities, and functional status; medication use and health care utilization (e.g., emergency room visits); and family knowledge about asthma management. These questions and related data collection, such as lung function tests, would be expected to be collected in a study focused on childhood asthma, but it is not clear how the NCS would be able to collect such detailed information on all domains of child health, development, and disease. The NCS documents mention the need to identify priorities for data collection, but the discussion does not adequately address how the NCS will overcome what is likely to be a major impediment in operationalizing the health phenotype concept, since respondent burden and limitations in collecting biological and environmental specimens will be critical considerations in designing the study content.

CONCLUSION 4-2: While using a dynamic health phenotype concept to plan the content of the National Children's Study appears to be a promising strategy, the panel lacked sufficient information to judge whether the implementation of such an approach would be feasible given constraints imposed by respondent burden and overall study costs.

Conceptual Framework for Measuring Health and Development

Although the use of domains and primary observations, rather than predetermined disease categories, does not necessarily require a shift from the original disease-oriented focus of the NCS, the current proposal tends to deemphasize a focus on disease outcomes and gives greater emphasis to positive health and development domains. The prior study (National Research Council and Institute of Medicine, 2008) expressed concern that there was no apparent overarching conceptual framework for health and development to tie the study together. In response, the current NCS proposal describes a detailed conceptualization

of health and development that consists of seven domains (demographics, physical health, psychosocial, neurodevelopmental, health behaviors, social environment, physical environment), each with subdomains (NICHD, 2013b, pp. 29-31). The NCS proposal indicates that these domains have been prioritized on the basis of gaps identified through literature review, public health significance, and the need for NCS platform and suitable instrumentation.

This approach has used the work of NICHD's existing Health Measurements Network, which views health as multidimensional with each dimension being assessed from very low to very high levels.[4] Each dimension includes several domains, and these can be assessed through multiple measurement modalities. Health is the product of a complex and dynamic interaction between the child and its environment (e.g., NICHD, 2013d, p. 27). These domains and subdomains can be considered along two axes, health and development. Dimensions for the health axis include adaptability, experience, function, and potential. For the development axis, the dimensions are plasticity, experience, and complexity. Although NCS proposes to look at a variety of dimensions of health, many concepts, such as functional status and severity of illness, remain unclear.

The panel judges that the conceptualization of health and development represents a substantial advance from the one reviewed in the previous evaluation (National Research Council and Institute of Medicine, 2008). The breadth of the conceptualization would encompass most of the issues affecting child health and development and provide many dimensions that could be linked to environmental exposures. The delineation of domains and subdomains is detailed enough to suggest quite specific measures that would need to be obtained.

However, as with the health phenotype concept, the panel did not receive needed details on the operationalization and effectiveness of the new conceptualization. Since no information of the specific measures for domains and subdomains was provided, the level of detail to be obtained is uncertain. While there may be time to develop measurement strategies for later years, the proposed 2015 start date for the NCS requires that the data collection rationale and strategy be more fully developed for at least pregnancy and the first year of life.

CONCLUSION 4-3: Many of the principles and concepts guiding development of the study design and the concept of having processes for developing future hypotheses and study content are consistent with the study platform framework for the National Children's Study. However, it is not clear whether and how those principles and concepts can be effectively used to design the study content.

[4]For a description, see: http://www.nationalchildrensstudy.gov/research/workshops/Pages/Forrest-Metadata-workshop-jan-2012.pdf [May 2014].

PROPOSED DATA COLLECTION

Study Visit Schedule and Mode

The panel was asked to comment on the more intense schedule early in the study compared with later years. Tables 4-1 and 4-2 show the proposed study visit schedule and selected content. For purposes of comparison, the tables also shows the same materials proposed for the previous review (National Research Council and Institute of Medicine, 2008). Table 4-1 displays the study visit schedule with mode and notes the planned collection of biospecimens, environmental measures, questionnaire content, and examinations from pregnancy through age 3. Table 4-2 shows the study visit schedule and collection mode planned for ages 3.5 through 21. The concentration of data collection in the early years is apparent in both the 2008 and 2013 schedules, with the current schedule including two prenatal interviews (rather than three as in the previous plan) and four data collections between 3 and 12 months of age.

Although there is reasonable scientific justification to conduct more frequent data collection during the prenatal period and early years, the documents provided to the panel do not explain adequately the scientific basis for the specific schedule of visits proposed for the NCS. For example, in view of concerns about study cost, it is not clear why the specific 3-, 6-, 9-, and 12-month schedule is needed in the first year.[5]

The panel requested that the NCS provide a rationale for the proposed study schedule. The first response (NICHD, 2013d, p. 75) was "Early and frequent data collection will help build health profiles as well as collect data during periods of rapid development. Operationally the time periods must be standard and easy to apply to a large cohort." A second response (NICHD, 2013h, p. 2) was

> The rationale for the proposed data collection intervals is based on the need for frequent data collection during periods of rapid change. The proposed visit schedule is intended to capture information about critical developmental events with the greatest precision. The data collection framework is based on a life course research model following extensive consultation with multiple stakeholders over a two-year period. Alternative schedules with less frequent visits were considered but rejected to avoid gaps in data collection opportunities during critical developmental periods. In addition, empiric experience supports frequent visits to increase retention and improve participant tracing.

[5]The cost model described in Chapter 5 shows that the incremental cost of a single in-person interview in the child's first year adds $90 million to the cost of fielding the study. The incremental field costs of a telephone interview total $45 million.

TABLE 4-1 Proposed NCS Study Visit Schedule: Comparison of Time, Mode, and Types of Measures Between the 2008 and Current Plans: Pregnancy to Age 3

Time of Measurement	Mode	Type of Measure: 2008 Plan[a]	Type of Measure: Current Plan[b]
Prepregnancy	Home visit	Biosamples: blood [M], urine [M], saliva [M], vaginal swabs [M], hair [M]	
		Environmental: indoor air, house dust	
		Questionnaire/diary: demographics, household composition, medication use, health behaviors, housing characteristics, chemical exposures, product use, occupational exposures, diet	
		Examinations: anthropometrics [M], blood pressure [M]	
Prepregnancy	Telephone	Questionnaire/diary: diet, chemical exposures	
First trimester	Home visit	Biosamples: blood [M, P], urine [M, P], saliva [M], **vaginal swabs [M], hair [M, P]**	Biosamples: blood [M, P], urine [M, P], saliva [M, **P**]
		Environmental: indoor air, house dust, drinking water, **soil**	Environmental: Air, dust, water, **visual assessment of house and neighborhood**

continued

TABLE 4-1 Continued

Time of Measurement	Mode	Type of Measure: 2008 Plan[a]	Type of Measure: Current Plan[b]
		Questionnaire/diary: demographics [M, P], household composition [M, P], medication use, health behaviors, housing characteristics, chemical exposures, product use, occupational exposures, diet, medical history [M, P], stress and social support [M], depression [M], tobacco use [P], cognition [P] (all updates if second interview)	Questionnaire/diary: [M, P]: demographics, medical conditions, disease exposure history, psychosocial, household occupational or hobbies exposure to chemicals, health behaviors, family medical history. [M]: medications, consumer product use, treatments and medical events, pain or other complications, dietary assessment
		Examinations: anthropometrics [M, P], blood pressure [M, P], fetal ultrasound (from med report or clinic visit)	Examinations: anthropometrics [M, P], blood pressure [M, P], fetal ultrasound (from med report or study administered)
Second trimester	Telephone	Questionnaire/diary: major life events [M], and updates [M] on mental health, medical, chemical exposures, and housing	
Third trimester	Clinic visit	Biosamples: blood [M], urine [M], saliva [M], vaginal swabs, **hair [M]**	Biosamples: blood [M], urine [M], saliva [M], vaginal swab
		Environmental: **indoor air**, house dust (self-collected and mailed in)	Environmental: house dust, **visual assessment of house and neighborhood**
		Questionnaire/diary: updates from [M] on demographics, household composition, medication use, health behaviors, housing characteristics, chemical exposures, product use, occupational exposures, diet, medical	Questionnaire/diary: psychosocial occupational and hobby exposures, health behaviors, medications, treatments and medical events, consumer product use, dietary assessment

		history, stress and social support, prenatal life events, depression.	
		Examinations: anthropometrics [M], blood pressure [M], fetal ultrasound	Examinations: anthropometrics [M], blood pressure [M], 2-D fetal ultrasound (subsample)
Birth	Hospital	Biosamples: blood [M], **urine [M]**, cord blood, placenta and cord samples, heal stick [C])	Biosamples: blood [M], cord blood, placental weight, photo and tissue, **meconium, infant skin, stool, and oral swab,** blood spot [C], **breast milk (1 month mail or pick-up)**
			Environmental: If enrolled at hospital (air, dust, water to be self-collected and mailed in)
		Questionnaire/diary: health behaviors [M], diet [M], chemical exposures [M], plans for infant feeding, sleeping, etc.	Questionnaire/diary: [M]: medications, consumer product use, treatments and medical events, pain or other complications, recent medical social and environmental history, planned health behaviors, medical record review including abstration for hearing screen and neonatal exam
		Examinations: anthropometrics [C], **dysmorphology and neurologic exam [C], digital photographs of face and anomalies [C], chart abstraction [M, C]**	Examinations: neonatal anthropometry

continued

TABLE 4-1 Continued

Time of Measurement	Mode	Type of Measure: 2008 Plan[a]	Type of Measure: Current Plan[b]
3 Months	Telephone/ remote	Biosamples: breast milk mailed in	Biosamples: breast milk mail or pick-up
		Questionnaire/diary: child care, medical update [C]	Questionnaire/diary: age-specific and other modules, dietary assessment, neurodevelopmental
6 Months	Home visit	Biosamples: urine [C], **hair [C], saliva [M, P], breast milk**	Biosamples: urine [**M,** C], **blood [M], skin, stool and oral swab [C]**
		Environmental: indoor air, house dust, drinking water, **soil,** visual assessment of house and neighborhood	Environmental: air, dust, water, visual assessment of house and neighborhood
		Questionnaire/diary: stress and social support, family process and parenting practices [M, P], health behaviors [M], depression and cognition [M], diet [C], medical update [C], medication use [C], media exposure [C], child care, chemical exposures, temperament [C], tobacco use [P], cognition [P]	Questionnaire/diary: [M]: core questionnaire, age-specific and other modules, dietary assessment, neurodevelopmental
		Examinations: anthromorphics [C], **dysmorphology exam and photos [C], dermatologic exam [C], social development observation [M, C]**	Examinations: infant anthropometry

9 Months	Telephone/ remote	Questionnaire/diary: child care, medical update [C], housing update, chemical and occupational exposures [M, C]	Biosamples: **breast milk mail or pick-up** Questionnaire/diary: age-specific and other modules, dietary assessment, neurodevelopmental
1 Year	Home visit	Biosamples: blood [C], urine [C], **hair [C], saliva [C], breast milk**	Biosamples: urine [**A**, C], blood [**A**, C], **skin, stool and oral swab [C]**
		Environmental: indoor air, house dust, drinking water, **soil,** visual assessment, noise survey	Environmental: air, dust, water, visual assessment of house and neighborhood
		Questionnaire/diary: household composition update, family process and parenting practices [M, P], health behaviors [M], diet [C], medical update [C], medication use [C], media exposure [C], child care, housing update, chemical and occupational exposures [M, C], language acquisition and social interaction [C], tobacco use, cognition (if not assessed at first trimester)	Questionnaire/diary: core questionnaire, age specific and other modules, dietary assessment, neurodevelopmental
		Examinations: anthromorphics [C], **blood pressure [C], dermatologic exam [C], cognitive exam [C], motor and language assessments [C], social development observation [P, C]**	Examinations: infant anthropometry

continued

TABLE 4-1 Continued

Time of Measurement	Mode	Type of Measure: 2008 Plan[a]	Type of Measure: Current Plan[b]
18 Months	Telephone/remote	Questionnaire/diary: child care, medical update [C], diet [C], housing update, chemical and occupational exposures [M, C]	Questionnaire/diary: core questionnaire, age specific and other modules, dietary assessment, neurodevelopmental
2 Years	Telephone	Environmental: indoor air and house dust self-collected and mailed in Questionnaire/diary: child care, medical update [C], housing update, chemical and occupational exposures [M, C], life events [M]	
	Home or other visit		Biosamples: urine [M, C], blood [M, C], skin, stool and oral swab [C] Environmental: air, dust, water, visual assessment of house and neighborhood Questionnaire/diary: core questionnaire, age specific and other modules, dietary assessment, neurodevelopmental Examinations: child anthropometry
2.5 years	Remote[c]		

3-years	Clinic visit	
	Home or other visit	Biosamples: urine [A, C], blood [A, C], saliva [A, C] Environmental: noise, dust, visual assessment of house and neighborhood. Questionnaire/diary: core questionnaire, noise exposures, risk and safety behaviors, social activities, physical activity, sun exposure, toilet training, occupational/ hobby exposures, reported height and weight, Ages and Stages, SWAN, NIH Toolbox Early Childhood Cognition Battery, Neuro-Psycholosocial Direct Observation Data Collector Instrument, WAST, Major Life Events, alcohol tobacco and substance use. Examinations: height or length, weight, circumferences, upper arm length, blood pressure, physical activity (subsample), vision screening.

NOTES: The entries in boldface type are survey characteristics that are different between the earlier and current plans. C, child; M, maternal; P, paternal; A, adult living with child, preferably first-degree relative. Mode: remote, telephone, internet, or mail.

[a]Information from National Research Council and Institute of Medicine (2008, pp. 26-31).

[b]Information from NICHD (2013b, pp. 42-47).

[c]There is no information on what data would be collected.

TABLE 4-2 Proposed NCS Study Visit Schedule: Comparison of Time and Mode for the 2008 and Current Plans, Ages 3.5 to 21

Time of Measurement	Mode: 2008 Plan[a]	Mode: Current Plan[b]
3.5 Years		Remote
4 Years		Home or other visit
4.5 Years		Remote
5 Years	In-home or clinic	Home or other visit
7 Years	In-home or clinic	Home or other visit
9 Years	In-home or clinic	Home or other visit
11 Years		Home or other visit
12 Years	In-home or clinic	
13 Years		Home or other visit
15 Years		Home or other visit
16 Years	In-home or clinic	
17 Years		Home or other visit
19 Years		Home or other visit
20 Years	In-home or clinic	
21 Years		Home or other visit

NOTES: Neither the 2008 nor current plan specifies what data would be collected at the visits.

[a]Information from National Research Council and Institute of Medicine (2008, pp. 26-31). The report (p. 32) states the schedule after age 5 is provisional, and there may be phone calls at more frequent intervals.

[b]Information from NICHD (2013b, pp. 42-47).

These statements provide a rationale for having a frequent visit schedule based on operational considerations; however, they do not describe the scientific basis for choosing the precise study visits schedule and content.

The NCS documents suggest that it is not possible to provide justification for the particular data collection schedule because the "the specific times of vulnerability [to environmental exposures that influence child growth and development] remain largely unknown" (NICHD, 2013d, p. 31). Although specific etiological windows for many exposures may be unclear, much is known of the human developmental process that could also guide the selection of critical periods for data collection. Furthermore, other aspects of health and development could be used to guide the assessment schedule, such as documentation of key developmental tasks; emergence of specific health or developmental problems; characterization of interventions that might improve outcomes; or even methodological issues, such as cohort maintenance. Besides issues of

persistent and transient environmental exposures, there is a robust research literature on factors influencing the duration of recall that might also guide the selection of assessment points. Thus, the response provided to the panel that some windows of vulnerability are unknown is an insufficient justification for the specific proposed study visit schedule.

CONCLUSION 4-4: The panel agrees that more intensive data collection in the early years of the National Children's Study is important, but the panel did not receive sufficient scientific justification to assess the merits of the specific data collection schedule.

Preconception Data Collection

As discussed in Chapter 2, a key goal of the earlier NCS research plan was to enroll nonpregnant females through the household-based sampling frame in order to collect data during the preconception period for some of the participants. This goal is retained in the current proposal, although the NCS now proposes to collect preconception data on subsequent siblings of enrolled participants by collecting data on the families and enrolling the subsequent siblings with certainty. In addition, it would supplement this group by enrolling a convenience sample of mostly nulliparous nonpregnant women for preconception data collection. As stated in Chapter 2, the panel acknowledges the potential scientific value of gathering preconception exposure data, but it recommends that a supplemental convenience sample not be included due to the very high recruitment and data collection costs associated with such a sample. (See Chapter 2 for additional discussion of a convenience sample.)

The current design would obtain preconception data through enrollment of subsequent siblings of enrolled participants. A potential benefit of this strategy is that it would minimize data collection cost because much of the data collection, such as collection of mother's biological specimens or a household dust specimen, are common to almost all of the in-person visits and would be relevant in assessing the preconception environment of a subsequent sibling. Nevertheless, it is not clear whether or how the planned data collection would be modified during the postnatal period for enrolled families to be able to address scientific hypotheses related to the preconception environment of the not-yet-born subsequent sibling. In addition, the current proposed data collection schedule for families when the target child is between 1 and 5 years of age, when the vast majority of subsequent siblings would be conceived, includes only one in-person and one remote data collection event per year. The intensity of the data collection visits may not be sufficient to be able assess the preconception environment, considering the need to adjust for varying times to conception.

Prenatal Data Collection

The scientific rationale for the NCS to enroll pregnant women and collect data during the prenatal period is detailed in Chapters 1 and 2. A substantial body of scientific research suggests that multiple social, biological, and physical factors during the prenatal period could affect child health and development and that the effects of these factors could vary during the pregnancy due to the developmental process. Although the prenatal period data collection protocol for the Main Study has not been finalized, the NCS documents list the most important domains and subdomains the NCS will try to measure.

The importance of collecting prenatal data derives in part from the fact that many factors to be measured, such as diet or medication use, may not be reliably recalled in postnatal questionnaires. Also, as indicated in the NCS documents (NICHD, 2013a, App. 4; 2013b, pp. 50-51), many environmental factors that could be measured in biological or environmental specimens are not persistent and might not be measured in specimens collected during or after the birth visit. Thus, prenatal data collection is essential to ensure concurrent collection of key social, biological, and environmental data.

RECOMMENDATION 4-3: The National Children's Study Main Study should collect data during the prenatal period at multiple times for as many of the study participants as the budget will allow.

Birth Enrollment and Data Collection

The current NCS proposal would enroll half of the probability sample (aside from subsequent siblings) at the time of birth in hospitals and birthing centers. Following the discussions in Chapter 2 on the scientific merit of this proposed strategy and in Chapter 3 on issues related to the sampling strategy, this section discusses issues related to the feasibility and quality of the data collection for women and children enrolled during a hospital admission at the time of the child's birth.

Enrolling and collecting high-quality data on women and babies in a hospital at the time of delivery presents several logistical challenges and concerns about the informed consent process. One NCS document (NICHD, 2013b, p. 23) indicates that on scheduled enrollment days, all women who are admitted to the hospital for possible deliveries would have to be screened for eligibility either by study staff or hospital personnel. Another document (NICHD, 2013g, p. 2) states that women could be approached for enrollment only after a minimum of 12 hours after delivery. The U. S. Office of Management and Budget, which must approve the protocols for all federal data collections, might require even a longer time interval.

Yet even the 12-hour minimum time interval could result in many potential participants being discharged before a recruitment visit and data collection.

Also because of the importance of collecting biological specimens, such as cord blood, as well as samples of the cord and placenta, the birth stratum enrollment protocol would necessitate making arrangements for these specimens to be collected on all potentially eligible women and babies prior to obtaining formal consent. In addition, the quality and validity of responses to a full baseline questionnaire administered postpartum to women in the hospital prior to their discharge may be less than optimal. It is also likely that the postpartum attrition of these participants would be greater than that of the women enrolled during the prenatal period who had already participated in an in-person data collection and agreed to continue with the study. In summary, there are many unclear and unresolved logistical issues related to the plan for a time-of-birth study enrollment and data collection. The NCS Vanguard Study has conducted only a very small pilot test of this enrollment and data collection strategy, which involved only two or three hospitals in each of three locations. The findings are still preliminary.

CONCLUSION 4-5: The strategy of the National Children's Study (NCS) to enroll a substantial proportion of participants at the time of the child's birth poses substantial logistical and operational challenges that have not been adequately tested in the NCS Vanguard Study.

RECOMMENDATION 4-4: Although the panel does not endorse the current proposal for a substantial birth enrollment stratum, if the National Children's Study (NCS) Main Study retains such a stratum, the NCS should conduct a full pilot test of recruitment and data collection during the birth visit before the Main Study is implemented.

Data Collection Content

The Program Office did not provide a detailed protocol or proposal for data collection, although the material it provided to the panel included examples of data and specimens that are being considered for inclusion in the protocol (see Table 4-1, above). The panel also did not receive draft questionnaires. The apparent lack of a draft final protocol and limited descriptions of possible data collection elements raises questions about the status of the NCS protocol. After several years of Vanguard Study pilot testing and based on the description of an elaborate process to develop study content, the panel expects that the NCS should be able to provide well-justified, near-final data collection protocols and study instruments, at least for the initial periods of the Main Study through the children's first year.

Study Visit Format

The NCS documents (e.g., NICHD, 2013b, p. 33) describe a strategy for the study visits to manage participant burden that consists of administering a core questionnaire to all participants at each visit and supplementing this questionnaire with modules on individual topics. The topics of the modules could address such factors as those related to the child's age at the time of the visit, a particular exposure, a new diagnosis, or a change in household composition. The documents indicate that the modules could be administered on the basis of contextual triggers or through random assignment (e.g., for validation or to collect information on a control group for which there was no contextual trigger). In addition to a questionnaire, the modules could include additional modalities for data capture, such as images or environmental specimens.

The proposed study format strategy seems to be well conceived and necessary in order to achieve reasonable respondent burden. The documents contained a draft list of domains to be addressed in a core questionnaire, but did not include a questionnaire. The documents also did not detail the process of selecting and prioritizing modules in real time prior to or during a data collection visit. It seems possible, for example, that participants who live in socially and environmentally disadvantaged homes and neighborhoods with poor access to schools, social services, and medical care and who have multiple health and development conditions could have a very large number of contextual triggers for the additional modules. Due to the lack of detailed information on how the NCS would implement the strategy of a core plus modules and the actual measures to be used, the panel cannot assess whether the proposed strategy would be able to contain respondent burden while collecting the data needed to characterize outcomes, identify key issues for health disparities, and operationalize the health phenotype concept.

Environmental Assessments

As noted in Chapter 1, the Children's Health Act of 2000 mandated that the NCS should be planned to be a "longitudinal observational birth cohort study to evaluate the effects of chronic and intermittent exposures on child health and human development in U. S. children" (P.L. 106-310). Environmental assessment—in which "environment" is broadly conceived to encompass social, biological, physical, chemical and other factors—is a critical component of the NCS study content. Unfortunately, high-quality environmental assessment can be expensive if it involves collecting media (e.g., dust, air, water) on multiple occasions and then processing, archiving, and analyzing the media for many possible agents. The NCS has to strike a balance between study cost and the imperative to collect a sufficient amount of high-quality data needed for environmental assessment.

The NCS documents (e.g., NICHD, 2013b, App. 4) describe the general

approach to environmental assessment and provide tables that list examples of biological or environmental specimens that could be used to measure different potential exposures. The information provided in these documents is not specific, and the list of specimens to be collected for study visits is considered to be preliminary. As previously mentioned, although the congressional mandate for this panel's study called for "a comprehensive review and issue a report regarding proposed methodologies for the NCS Main Study," the NICHD did not ask the panel to review the environmental assessments, and the panel did not receive sufficient information to evaluate the scientific merit of the draft environmental assessment protocols.

When discussing the environmental assessment, the NCS documents (e.g., NICHD, 2013b, p. 48) refer to National Research Council and Institute of Medicine (2013), a summary of a workshop on the design of the NCS, and an earlier workshop convened jointly by the NCS, the Environmental Protection Agency, and the National Institute of Environmental Health Sciences in 2010 to review the specific exposure matrices for the NCS. The NCS documents quote statements from the workshop speakers to indicate there was a consensus that the NCS strategy and plans for environmental assessments were reasonable. However, the panel's review of the workshops' reports (U.S. Environmental Protection Agency, 2010; National Research Council and Institute of Medicine, 2013) found that important caveats and critical measurement issues discussed in the two workshops were not sufficiently acknowledged by the NCS in the documents provided to the panel.

The exposure assessment protocols reviewed and discussed by the workshop participants were based on earlier protocols that included collection of biological specimens and "air, dust, water" during multiple study visits. The documents provided to the panel did not have the same extensiveness of biological or environmental data collection as was being considered in 2010. The findings of the workshops are not necessarily applicable to the environmental assessment currently being considered. Furthermore, the workshop participants expressed concerns about the validity of using questionnaires to assess many types of environmental factors and especially to rely on retrospective recall of exposures. For example, the primary exposure source of many non-persistent hormonally active agents, such as phthalates and bisphenol A, is consumer product use. The report of the earlier workshop (U.S. Environmental Protection Agency, 2010, p. 6) notes: "However, most adult participants cannot provide sufficiently accurate information for classifying exposures based on product use or activities." The workshop participants noted that these agents or their metabolites can be measured in biological specimens, but it would be important to collect specimens during multiple data collection visits because of the short half-life of the metabolites in urine. A similar comment about using questionnaires to assess chemical exposures was made in the second workshop (National Research Council and Institute of Medicine, 2013, p. 22):

"For example, you can't ask people if they have PCBs in their home or if they have polybrominated diphenyl ether flame retardants in their TVs or couches."

The exposure assessment experts at the two workshops emphasized the importance of collecting biological and environmental specimens during multiple data collection visits during the critical time periods of development. The summary of the 2010 workshop stated (U.S. Environmental Protection Agency, 2010, p. 2)

> All workgroups agreed that in utero and through early childhood (up to ages 3 to 5 years) were the time periods when children were most susceptible and when exposure monitoring should be conducted. At a minimum, all groups preferred to conduct monitoring during three visits, one each during the first trimester, the third trimester, and the first year.

Retrospective exposure assessment based on bulk dust samples may be used to assess average exposures to persistent metals and chemicals over a period of several months, but it is not a viable strategy to assess transient exposures to nonpersistent agents. Furthermore, the two workshops did not endorse the current NCS proposal to conduct a retrospective exposure assessment for families enrolled at the time of the child's birth by providing collection kits to the families for self-collection of environmental samples. The panel also judges that the proposed methods to assess environmental exposures by relying on maternal collection of in-home environmental samples have not been adequately pilot tested.

CONCLUSION 4-6: Exposure assessment, including collection of biological and environmental specimens during multiple study visits beginning during the prenatal period, is a critical component of the National Children's Study in addressing the mandate of the Children's Health Act of 2000 and fulfilling the study's goal to serve as a platform for future scientific inquiry.

Process for Selection of Measures

Specific measures for the measurement domains and subdomains have not yet been specified. Although the NCS provided a detailed description of the advisory and consultative process (NICHD, 2013b, pp. 34-35; 2013g, pp. 7-8) to inform the decision making for measurement methods and instruments, it is not clear how the advisory and consultative process actually informs decision making. The process is extensive, but seemingly unwieldy for timely development of protocols and study instruments. The panel received no documentation that the process for developing measurement methods and instruments has been formally evaluated or compared with other large national and international longitudinal cohort studies. Nor did it receive specific documentation or

evidence of the process in action with any domains, subdomains, or instruments and how it resulted in instruments or testing in the Vanguard Study.

Considering that Vanguard Study field work started in 2009 with more than 4,000 families enrolled and followed at least through the birth visit and that the advisory and consultative process has been in place since at least that time, it would seem that the NCS ought to be able to provide documentation of nearly final data collection methods for the prenatal period and child visits through 6 months of age based on this experience. The comments by the Program Office that findings and data from the Vanguard Study pilot studies are still being evaluated, and reports from the NCS Federal Advisory Committee that the Committee also has not yet been provided draft data collection instruments and methods for review, reinforce this concern.

CONCLUSION 4-7: The processes for developing content for the National Children's Study are complicated, and insufficient documentation has been provided to demonstrate that the processes will be effective.

RECOMMENDATION 4-5: The National Children's Study Program Office should document and provide justification for development of the data collection schedule, content, and methods now and going forward. The documentation should be sufficient to guide use of the study data by future researchers.

RECOMMENDATION 4-6: The National Children's Study Program Office should finalize the study visit data collection protocols that it intends to use for the Main Study (including questionnaires and other measurements), at least through age 1, and then pilot test the protocols before implementing the Main Study. The protocols and findings of the pilot tests should be peer reviewed and approved by the proposed independent oversight committee prior to initiating the Main Study.

(See Chapter 6 for discussion and recommendations regarding an oversight committee.)

HEALTH DISPARITIES

As discussed in Chapter 2, the NCS proposes to address health disparities during data collection by ensuring that information about demographic and other characteristics that define these populations is gathered in the core questionnaire and measuring exposures that may be important for understanding health disparities.

Based on the responses to panel questions, the NCS clarified that the major domains of interest for health disparities are race and ethnicity, socioeconomic

status, geography, and immigration status, stating that the NCS will follow the Data Collection Standards of the U.S. Department of Health and Human Services[6] for collecting information on race, ethnicity, sex, primary language, and disability status. Some questions will be included to assess immigration status, and in addition, information on health insurance status, other health care access characteristics, and education will be collected. The document also noted (NICHD, 2013d, p. 79) that geography can be used to investigate urban and rural differences, and also to identify specific industrial exposures common in some areas. Although the domains identified by the NCS documents are standard and reasonable, there was no indication that the NCS has developed or adopted a conceptual framework for health disparities (e.g., similar to the framework the NCS has developed to guide assessment of child development) or a strategy to identify additional domains and measures relevant to health disparities, such as psychosocial factors or features of social or physical environments that may be of special relevance to understanding health disparities.

CONCLUSION 4-8: Based on the information provided, the panel concludes that the National Children's Study plan has paid insufficient attention to how health disparities should be taken into account in the development of the schedule visit and content of the Main Study.

RECOMMENDATION 4-7: The relevance to health disparities should be an explicit criterion for selecting the constructs that will be assessed as part of the National Children's Main Study, the measures that will be used to assess them, and the timing of the assessments. The NCS should obtain input from experts on health disparities in childhood as part of the documented process through which the measures for inclusion are selected, and the measures should be approved by the proposed oversight committee.

(See Chapter 6 for discussion and recommendations regarding an oversight committee.)

DATA RELEASE

As noted in the previous review (National Research Council and Institute of Medicine, 2008, p. 199): "Past experience with virtually all national data sets is that the research value of the data is maximized when as many skilled analysts as possible are able to access the data for original and replication analyses, and when the peer-review process judges the quality of the analyses performed." It recommended: "[T]he NCS should begin planning for the rapid dissemination

[6]Available at: http://aspe.hhs.gov/datacncl/standards/aca/4302/index.pdf [April 2014].

of the core study data, subject to respondent protection, to the general research community."

Guttmacher et al. (2013, pp. 1873-1874) describe a reassuringly open data release policy for the study:

> The NCS is committed to broad, rapid sharing of all data and samples, while respecting participants' privacy and confidentiality. No individuals or institutions that gather data and samples will have prioritized claims to them. Electronic data will be available to all qualified researchers through controlled access mechanisms, in keeping with current National Institutes of Health practices. Because biologic and environmental samples are exhaustible, there will be an application process for obtaining them. To maximize their use, the NCS will share promptly with the entire research community the results of all analyses performed.

Additional details about the NCS study data release policies were provided in documents made available to the panel (NICHD, 2013a, 2013d) and are based on review of the data release policies of a number of federal government and university-based surveys. The NCS expects a 2-year lag between the end of data collection and data release to the research community. It plans to release three types of analytic files. One would be a public-use file that would be "disseminated into the public domain without restrictions on access or use"[7] (NICHD, 2013d, p. 83). In order to protect confidentiality, in the public-use data, individual level data would be coded, aggregated, or otherwise altered to mask individually identified information" (p. 83). Use of the second type of data would be restricted through controlled access and use "through a licensing process whereby each data request is individually evaluated and, if approved, the data user enters into a formal data sharing agreement . . . [and] the approved environment for access . . . could include a 'Census-Bureau-type' data center" (p. 83). The third type of data would be controlled-use materials, such as environmental samples, biospecimens, images, and audio files. Access to these data or specimens would be even further restricted and would have to be approved because of the limited amount of specimens or because such data as images cannot be de-identified.

In order to develop and implement plans for data sharing, the NCS established the NCS Data Access Committee in 2009, which defined governing principles for data access and confidentiality. It also hosted a data use workshop in February 2013 with invitees from federal agencies, contract research organization, study centers, and other stakeholders (see NICHD, 2013a). Following the workshop, the NCS published a document on data access and confidentiality concept of operations (NICHD, 2013k), which provides more detailed information on the NCS data dissemination strategies.

[7]Public-use files are available for international and commercial use.

In its consideration of the NCS plans, the panel investigated the data release practices of a number of the studies referenced by NICHD (2013d). Perhaps the closest models for the NCS are the surveys conducted by the National Center for Health Statistics (NCHS) in the Centers for Disease Control and Prevention, which face the same general set of federal government constraints on data release as the National Children's Study.

The NCHS study that shares the most features with the NCS is the NCHS's National Health and Nutrition Examination Survey (NHANES). Although each cohort of NHANES is much smaller than the NCS (5,000 persons of all ages are interviewed each year) and is a repeated cross-sectional rather than longitudinal design,[8] it does involve personal interviews and collects data from physical examinations and laboratory tests, and it maintains a DNA repository on its samples. It also monitors environmental exposures and children's growth and development.

The general principles guiding data release for the NHANES are to distribute the data as widely as practicable, as soon as possible after data collection, and in as much detail as possible while maintaining survey participant confidentiality.[9] The NHANES data release performance matches these goals well. Almost all of the person-level survey, laboratory, and environmental data collected in NHANES are available to the public on the study's Website. The data are processed in 2-year cycles with the first data releases available within 9 months of the end of a given cycle's data collection period. Files for NHANES components that require longer to process are released as the datasets become available.

Confidential data, including DNA, imaging data, and geographic location, are made available to researchers under restricted data agreements. These data are made available through the NCHS Research Data Center (both on site and remotely) and through the Census Bureau's national network of Remote Data Centers.[10]

Given the similarities between NHANES and the National Children's Study, the panel views the general structure of the NHANES's data release policy and performance of NHANES as a model for the NCS. Confidentiality concerns arising from the longitudinal nature of the NCS may affect somewhat the balance of data released publicly and confidentially, but the panel would expect these kinds of changes to be relatively minor.

[8]The NHANES provides for the possibility of longitudinal follow-ups for its sample but does not routinely conduct such follow-ups.

[9]For details, see http://www.cdc.gov/nchs/data/nhanes/nhanes_release_policy.pdf [March 2014].

[10]For the NCHS center, see http://www.cdc.gov/rdc/ [March 2014]; for a description of the Census Bureau's centers, see http://www.census.gov/ces/rdcresearch/ [March 2014].

CONCLUSION 4-9: The panel endorses the general structure of the data distribution plans for the National Children's Study (NCS), although it fails to understand the need for a 2-year lag between the availability of analytic data and their release to the research community. Subject to confidentiality concerns, timely and complete data access are vital to maximize the scientific value of the NCS and have been achieved by other federal government surveys, which ought to serve as models for the NCS.

Finally, the panel considered another challenging issue related to data release. Given the nature of the recruitment cycle and the roll out of the survey at different times in different primary sampling units (PSUs), with the 4-year roll-out period assumed in the panel's cost analysis, it will be 7 years before any given data item (e.g., from the questionnaire administered during the 1-year visit) has been collected for all children enrolled in the study. The reason, as explained in Chapter 3, is that the PSUs will be divided into groups, so the field work will be implemented in one group each in 4 successive annual years, called waves, with the birth window being 4 years in each location. Therefore, it will take 7 years from the first data collection in the first wave of PSUs to the last data collection in the last wave of PSUs. Given this lengthy interval, it is imperative to develop data processing and the documentation associated with data release based on data gathered in the first few years, so that minimal effort will be needed to release a given wave's data after its data have been collected.

Beyond a rapid end-of-wave data release, the 7-year data cycle argues for consideration of an "early release" data policy 2 or 3 years into the cycle to encourage data quality exploration. Given the complications of the sampling design, these preliminary data could not be used to generate national or local estimates. But if experienced analysts were provided access to these data through the proposed network of "restricted access" data centers, a great deal could be learned about properties and quality of these data, in particular, newly developed interview and observational data. This approach would improve the quality and timely release of complete-wave data and their documentation, and it would likely inform the design of recurrent question and observation sequences in future waves. It will be important for the NCS to clearly state to prospective analysts that such data are incomplete and not representative. An "early release" policy would increase processing costs somewhat, so the value of the policy would need to be judged against its costs.

RECOMMENDATION 4-8: The panel recommends that the National Children's Study should consider producing an "early release" version of the data from the Main Study that includes data collected in the early years of each wave's data collection cycle and makes those data available to analysts under the terms of restricted access data centers.

5

Data Collection Costs

The panel sought to ground its recommendations for the design of the National Children's Study (NCS) in an understanding of the nature of likely field costs and by calculating the field costs of alternative sample designs and field strategies over the period of respondent recruitment. Although this topic was not explicitly included in the charge to the panel from the National Institute of Child Health and Human Development (NICHD), the panel judged that an evaluation of the scientific merits of the proposed methodologies of the NCS Main Study required an understanding of the relative benefits and costs of design options. It is important to note that the cost figures we have developed are based only on major costs of data and specimen collection and thus understate, perhaps substantially, total project costs over the period we considered. (Examples of the omitted cost categories are provided in Appendix B.)

ASSUMPTIONS

In the course of developing our estimates, we provided the NCS Program Office with a list of the major assumptions used in our cost model (see Appendix B for a more complete description of our assumptions). The Program Office agreed with substantially all of our cost assumptions, and we accepted most of the changes to our assumptions and estimates (either up or down) suggested by Program Office staff. One noteworthy exception concerned what we believe is

the Program Office's likely underestimation of the costs of persuading hospitals and providers to cooperate with the study.[1]

The previous chapters have identified three design issues with important cost implications: (1) the fraction of the main probability sample that would be recruited at the offices of prenatal providers affiliated with targeted hospitals (rather than being recruited when they presented for delivery); (2) the inclusion of a supplemental sample of nulliparous women to be recruited prior to the conception of their first births; and (3) the number of in-home and telephone interviews associated with each study participant. We modeled a total of eight different designs involving these dimensions.

In estimating field costs associated with alternative designs, our model only includes field operation tasks for which these issues would have significant cost effects. As detailed in Appendix B, it does not attempt to account for the differential costs associated with overall management at the program or contractor level, which include purchasing or building the sample frame, managing the sample and preparing sample weights, programming questionnaires, tracking and storing environmental or biological specimens, data entry, verification, transmission and management, archiving, documentation, and dissemination. In addition, because the NCS plan proposes that the sample of births be drawn over 4 years for each of four groups of primary sampling units, with the start-up of each birth window for each of these groups spread over 4 consecutive years, the enrollment phase will extend for 7 years. Thus, we carried out our cost model for 7 years, since the cost consequences of most of the design issues we investigate will have occurred during that period.

COST DRIVERS

Given the magnitude of the estimated 7-year total field costs of all of the scenarios we investigated—$1.3 to $1.6 billion—we begin by noting the most sizable components of field costs. We present these cost drivers in decreasing order of significance.

1. **Sample Size:** Pursuing a cohort of 100,000 is the primary driver of field costs. Even modest changes in this figure will generate substantial

[1]As detailed in Appendix B, the current NCS experience is based on the Vanguard Study. We judged that the involvement of academic medical centers in the Vanguard Study brought considerable name recognition to the recruitment effort. The kinds of national field organizations that are proposed for the Main Study are less likely to bring the same sort of name recognition and are likely to need more time, effort, and expense to secure the cooperation of hospitals and providers, and, possibly, even respondents. Although it is not possible to estimate these additional costs with much precision, additional outreach efforts, incentive payments, and more time on site to recruit respondents could translate into an additional $100 million in recruiting costs over the 7-year period (see Appendix B for the details on our assumptions).

dollar changes. A corollary to this fact is that if large cost adjustments are necessary, sample size is a necessary dimension in which to make large changes in projected costs.

2. **Recruitment Costs:** The cost of recruiting the baseline sample of 100,000 births is roughly $231 million or about 15 percent of the 7-year estimated field costs of the project. The total includes the costs of recruiting hospitals and providers, as well as the costs of recruiting the women whose births will become members of the probability sample. This cost varies somewhat with the fraction of the sample enrolled at delivery rather than prenatally.
3. **Interviewing Costs:** At roughly $1,000 each in current dollars, adding or dropping one face-to-face interview for each respondent changes 7-year total field costs by about $100 million. Adding or dropping one telephone interview per respondent changes 7-year total field costs by roughly $40 million.
4. **A Supplemental Sample of Nulliparous Women:** Because several women (about 3.6 based on current NCS estimates) need to be enrolled and given an in-home interview in order to yield 1 woman with a birth that will enter the sample, the cost of obtaining 5,000 preconception first births enrolled in the study is quite large, about $76 million.
5. **Fraction Recruited Prenatally:** Recruiting respondents prenatally and administering interviews prior to birth, rather than recruiting them at birth, increases project costs by $30 million for every 10 percentage point increase in the fraction of the total sample that is recruited prenatally.
6. **Interview Length:** Increasing by 5 minutes the length of all of the 1.05 million interviews administered in the first 7 years of the study, including both face-to-face and telephone interviews, increases total field costs by only $10 million. Attempting to save money by reducing interview length is thus relatively ineffective when compared with the other cost drivers.

FIELD COST ESTIMATES: ALTERNATIVE MODELS

Chapters 2 and 4 provide the scientific basis for recruiting into the probability sample as many women as possible prior to birth. Chapter 2 questioned the high cost and the lack of evidence of value of the 5,000-birth supplemental sample of preconception nulliparous women and the scientific rationale of supplemental samples of 5,000 that might be used for special purposes. Thus, the panel asked its cost consultants to explore a variety of approaches to maximizing prenatal recruitment while staying within the field cost constraints associated with the current design proposed for the NCS. This section presents the data on eight models that incorporate different assumptions about the extent of

prenatal recruiting, the inclusion of the supplemental samples, and reductions in the total number of interviews.

- **Model 1–baseline (current design):** 100,000 total births, consisting of 90,000 recruited as part of the probability sample (including subsequent siblings) and 10,000 recruited into the various supplemental samples. Of the 90,000 in the probability sample, the mothers of roughly 45,000 would be enrolled at their prenatal providers and receive one or two prenatal interviews. The remaining 45,000 would be recruited at the time of birth. Of the 10,000 supplemental-sample births, the mothers of 5,000 first births would be recruited prior to conception and the mothers of another 5,000 would be recruited at birth for the supplemental convenience sample.
- **Model 2–maximum prenatal recruitment:** This model isolates the additional expenses associated with maximizing prenatal recruitment in the current design. It is the same as the baseline except that 97 percent of the births recruited in the probability sample of 90,000 (including siblings) and the supplemental convenience sample receive prenatal interviews, along with the preconception sample of first births. We assume that 3 percent of births cannot be recruited prenatally and would be enrolled at delivery (see Chapter 2).
- **Model 3–maximum prenatal recruitment into the probability sample but no preconception first-birth supplemental sample:** This model isolates the incremental cost of the preconception supplemental sample. It assumes there is no supplemental sample of 5,000 preconception first births, and the probability sample is increased by 5,000 to 95,000; it retains the 5,000 in the convenience sample and assumes they are recruited prenatally. As in Model 2, 97 percent of births in the probability sample (and their subsequently enrolled siblings) are enrolled prenatally and the remaining 3 percent are enrolled at delivery.
- **Model 4–drop one in-home and one telephone interview:** This model and the next two make small changes in the number or types of interviews. This model adopts the assumptions in Model 3 and drops one face-to-face and one telephone interview in the child's first year of life.
- **Model 5–drop one telephone interview:** This model also adopts the assumptions in Model 3 and drops one telephone interview in the child's first year of life.
- **Model 6–drop one face-to-face interview:** This model also adopts the assumptions in Model 3 and drops one face-to-face interview in the child's first year of life.
- **Model 7–have only one prenatal interview:** This model also adopts the assumptions in Model 3 and includes only one interview prior to

birth for the 97 percent of the 95,000 in the probability sample that are recruited prenatally and the 5,000 in the convenience sample.

- **Model 8–maximum prenatal recruitment, no supplemental samples, and cost neutrality:** This model is similar to Model 3, but it incorporates maximum prenatal recruitment and cuts the number of sampled children to maintain cost neutrality with the current design. Its cost assumptions are the same as the baseline except that 97 percent of births in the probability sample (including siblings) are enrolled prenatally, and there is no preconception first-birth cohort or convenience sample. Under this model, 96,000 children could be enrolled in the study.

Table 5-1 shows the costs and number of births that could be included under these eight models. The second column shows the 7-year field costs, in millions, associated with these eight models. So, for example, the first row shows that the estimated field costs for the current study design total $1.495 billion over 7 years. The third column shows the incremental field savings (or costs) associated with Models 2-8 relative to the baseline cost projection for the field work. The fourth column shows, for each model, the size of the sample that would be needed to make that field strategy cost neutral with the baseline model. The central message of Table 5-1 is that, although the costs of administering the prenatal protocol to women recruited at their providers are higher than the costs of recruiting them at delivery, by eliminating the preconception cohorts for nulliparous women (but retaining preconception cohorts in the sibling samples), the NCS can attain cost neutrality either with modest changes to the data collection protocol or by modifying the sample size. If, as recommended in Chapter 2, the other 5,000 supplemental sample component of the proposed design is dropped, then, at cost neutrality, the size of the probability sample is either close to or larger than what is achievable with the baseline specification.

CONCLUSION 5-1: The panel estimates that the field costs of the current plan for the National Children's Study over its first 7 years will total $1.495 billion. This estimate does not include a number of other sizable contributors to overall study costs.

Turning from 7-year total field costs to the associated pattern of annual field costs, Figure 5-1 shows a linear increase in annual field costs from about $60 million in the first year to more than $300 million in the fourth year of the study. For reference, the fiscal 2010-2012 congressional appropriations for the NCS were each around $190 million; the fiscal 2013 and 2014 appropriations

TABLE 5-1 Field Costs of Alternative Sample Designs

Model and Sample Size*	Total Field Cost (in millions of $)	Base Field Cost–Modeled Field Cost (in millions of $)	Size of Main Sample (including siblings) for Field Cost Neutrality
N = 100,000			
Baseline: current design—one-half prenatal and one-half birth recruitment for 90,000 births, plus 5,000 preconception supplemental sample and 5,000 convenience sample (Model 1)	$1,495	NA	NA
Maximum prenatal recruitment plus preconception supplemental sample, and convenience sample (Model 2)	$1,631	($135)	89,179
Maximum prenatal recruitment and convenience sample but no preconception supplemental sample (Model 3)	$1,542	($47)	96,256
Drop one in-home and one telephone interview from Model 3 (Model 4)	$1,348	$147	113,940
Drop one telephone interview from Model 3 (Model 5)	$1,492	$3	100,256
Drop one in-home interview from Model 3 (Model 6)	$1,398	$98	108,836
Have only one prenatal interview in Model 3 (Model 7)	$1,488	$7	100,600
N = 96,000			
Maximum prenatal recruitment, no preconception or other supplemental sample and cost neutrality (Model 8)	$1,495	($0)	96,256

NOTES: All cost figures are based only on costs of fielding the study and thus do not represent the full cost of the NCS; see text for discussion. Cost of convenience sample assumed to be equal to cost of equal-sized prenatal sample.

*See text for details of the panel's models; see Appendix B for details of cost assumptions and other information about the panel's cost modeling exercise.

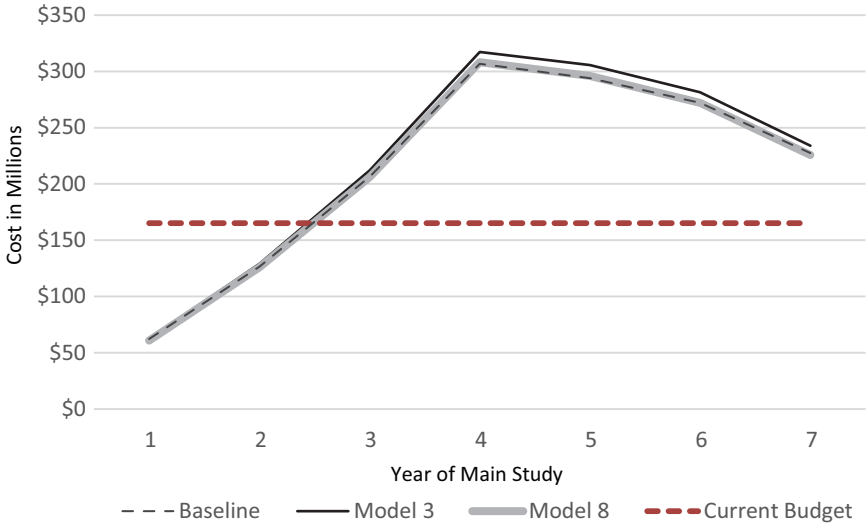

FIGURE 5-1 Annual field costs for three models of the NCS sample design.
NOTES: The costs include only field costs. See text for discussion.
Model 1(baseline): NCS program office proposed plan (N = 100,000)
Model 3: Maximum prenatal recruitment but no preconception supplement sample (N = 100,000)
Model 8: Maximum prenatal recruitment, no preconception supplemental sample, and cost neutral (N = 96,000)

were up to $165 million,[2] respectively. It is important to remember that these are appropriations for total project costs and not just the field costs that we modeled.

EVALUATION

It is rarely, if ever, possible to accommodate all desired elements in the design of a single study, no matter how large, so that tradeoffs between cost and coverage always have to be considered before finalizing a design. An evaluation of those tradeoffs has to rest on an understanding of the likely benefits and costs

[2]Section 1508 of the Consolidated and Further Continuing Appropriations Act, 2013 states: "That $165,000,000 shall be for the National Children's Study (NCS), except that not later than July 15, 2013 the Director [of the NIH] shall estimate the amount needed for the NCS during fiscal year 2013, taking into account the succeeding proviso, and any funds in excess of the estimated need shall be transferred to and merged with the accounts for the various Institutes and Centers of NIH in proportion to their shares of total NIH appropriations made by this Act." The panel uses the term "up to" to describe this type of appropriation.

associated with the options under consideration. The panel was surprised to learn that the NCS does not appear to have developed a cost model for considering the likely tradeoffs. Consequently, the panel constructed the models above of some of the most important recruitment and interviewing costs likely to be incurred during the first 7 years of the Main Study. We note again that these components represent only a fraction of the total costs associated with the NCS.

As discussed above and in Table 5-1 and Figure 5-1 (above), the panel estimates total field costs of the current NCS design at about $1.5 billion, with annual costs ranging from $60 million in the first year to a little over $300 million in the fourth year. The average annual field cost is $214 million, significantly more than the most recent congressional appropriation for fiscal 2014 of up to $165 million. And that appropriation is not only for field costs, but for all costs of the NCS.

CONCLUSION 5-2: The major drivers of field costs for the National Children's Study Main Study are sample size and number of interviews. In-person interviews cost more than twice as much as telephone interviews. In contrast to contacting and gaining the cooperation of respondents, modest changes in interview length contribute minimally to cost.

CONCLUSION 5-3: For the same field costs and with the elimination of the 10,000 supplemental samples—5,000 nulliparous women and a 5,000 convenience sample—the National Children's Study could afford to enroll a predominantly prenatal probability sample of 96,000 cases with no other changes to the proposed data collection protocol.

It is to be hoped that adequate funding will be provided to carry out the proposed NCS Main Study design, after it has been refined and reviewed. But if adequate funding is not provided, tradeoffs will be required.

RECOMMENDATION 5-1: Given the goal for the National Children's Study (NCS) to understand the links of environmental exposures to child health and development and its cost structure, if major reductions in the cost of the study need to be made, they should be reductions in sample size rather than exposure domains. Along with such a decision to reduce the sample size, the NCS should reconsider whether to oversample minorities in order to maintain the ability to evaluate health disparities with a reduced sample.

6

Study Leadership and Scientific Oversight

This chapter addresses issues related to the leadership and oversight of the National Children's Study (NCS). Although the charge to the panel from the National Institute of Child Health and Human Development (NICHD) does not list this topic explicitly, the panel determined that it is a critical issue underlying the centerpiece of the panel's charge, which is the design of the Main Study and whether it "will produce scientifically sound results."

AREAS OF CONCERN

The legislation authorizing the NCS assigns responsibility for the conduct of the study to the NICHD (Section 1004a). It also requires (Section 1004b) the director of NICHD to establish a consortium of representatives from appropriate federal agencies (set up as the Interagency Coordinating Committee), including the Centers for Disease Control and Prevention (CDC) and the U.S. Environmental Protection Agency (EPA), to "plan, develop, and implement a prospective cohort study . . . ." While the authorizing language is ambiguous regarding whether NICHD or the federal agency consortium is "in charge" of the NCS, the NICHD Website states that the Office of the Director (which includes the NCS Program Office) has primary responsibility for planning and coordinating the NCS, and the NCS Website states that the NCS Program Office is responsible for the day-to-day operation of the NCS. The director of the NCS Program Office reports to the NICHD director and the director of the National Institutes of Health (NIH) on high-level design decisions. Beginning in 2010, with the decision that the previously approved sample design should be

changed, the NCS Program Office has been engaged in activities that are best described as study design and planning in addition to the ongoing operation of the Vanguard Study.

To provide input to the NCS, there is not only the Interagency Coordinating Committee, but also an NCS Federal Advisory Committee, which is chartered to advise the directors of NIH, NICHD, and the NCS Program Office. These two committees meet quarterly. For the Vanguard Study, the Program Office established a separate steering committee that included Program Office staff and field contractors, including university researchers. As the Vanguard Study was expanded, the steering committee was enlarged, and an executive steering committee was formed. The Program Office has also contracted for scientific expertise in particular areas, such as sample design and questionnaire content, through regular contract mechanisms, as well as through the establishment of various working group and committees. In recent years, it has provided publicly only summaries from this contracted work.

Problems with Decision Processes and Documentation

Despite these oversight and advisory structures, the panel finds that the processes by which study decisions are made and vetted are opaque. Moreover, decisions in a number of important instances are not well documented and do not appear to reflect the breadth or depth of relevant scientific expertise even though the NCS has engaged many well-qualified scientists since it began.

More broadly, it is the panel's view that the current management structure, with the Program Office in overall control of decision making for the NCS, is not likely to produce the optimal design for a study that needs to be implemented in a scientifically grounded and cost-effective manner. The panel's assessment on this issue is based on observations and experiences of panel members during its work and the preceding workshop (see National Research Council and Institute of Medicine, 2013); the panel's evaluation of the quality and completeness of the documentation provided to it (see Appendix A); and the experiences of panel members with the work of the Program Office over the course of the 6 years since the publication of the prior review of the National Children's Study (National Research Council and Institute of Medicine, 2008).

The panel found several troubling problems, including:

- proposed changes in sample designs with inadequate evidence on the quality of the proposed sample frames and lack of comparative analysis of the costs and benefits of the proposed changes;
- the absence of concrete plans for the supplemental sample components, the most costly of which would be the preconception sample of 5,000 nulliparous women;

- the apparent absence of a well-specified cost model for evaluating tradeoffs across alternative designs, an example of which is an inability to estimate the cost of crucial design components, such as prenatal as opposed to birth recruitment;
- insufficient detail concerning findings from the Vanguard Study[1] and incomplete study design plans, even though the Vanguard Study has been pilot-testing study design options and data collection protocols since 2009;
- lack of scientific documentation and specifications to support the proposed sample design, content, and other key decisions for the Main Study; and
- absence of a realistic and well-specified schedule of stages and decision points that demonstrate how and when a final, well-documented design for the Main Study would be complete and field implementation could begin. At the time of implementation, instrumentation should be available particularly for the early phases up to 1 year of age.

Need for Added Expertise

The panel sincerely appreciates the positive tone of interactions with the Program Office staff and with their expressions of willingness to provide responses to the panel's questions. However, as noted above and detailed in Appendix A, many of the responses were not fully responsive to the panel's requests. Thus, the panel is concerned that the Program Office may not have sufficient in-house expertise in relevant scientific and survey research disciplines[2] to enable it to function effectively in using the input it receives from contractors and advisory groups for design and operational decisions for the NCS Main Study.

Based on the material sent to the panel, the panel concludes that ongoing expertise is needed in sampling and weighting in longitudinal surveys, not only to draw the original sample, but also to deal with issues of attrition and missing data. In addition, expertise is needed in the design and management of

[1]The Program Office stated that" it plans to disseminate Vanguard data and is currently expecting to make its first release in late 2014" (NICHD 2013d, p. 86) and also stated that "many analyses are still pending, but early results have informed the sampling and recruitment design" (NICHD 2013g, p. 7). It also stated that the "Vanguard survey data have been released to 28 writing teams and 4 supplemental methodological survey projects for analysis for primary publication" (NICHD 2013f, p. 6).

[2]As noted in Chapter 1, the NCS Program Office has a staff of 18 who have advanced degrees in a variety of relevant disciplines. They are responsible for management and coordination of many contracts and advisory bodies. Although the current staff may have some of the areas of expertise recommended by the panel, the extent of their expertise was not evident in the documentation received by the panel. Moreover, the current staff are not likely to have time to conduct or stay abreast of research in relevant areas.

epidemiologic research in pregnant women and children in large, longitudinal studies. Expertise and experience are also needed in the planning and conduct of pilot studies, the management of multiple contracts, the development and maintenance of a sample collection system, and the timely production of analyzable datasets. Such expertise would require familiarity with the multiple disciplines required for this effort and how to deploy them efficiently. In particular, considering the NCS goal to assess the role of environmental factors on child health and development, the Program Office needs to have significant expertise in exposure assessment, expertise that encompasses a broad conceptualization of "environment."

Expertise in measurement theory and development is likely also needed. The NCS will face the need to have succinct, valid, and reliable measures throughout its course. For some issues, the measures can be developed with existing instruments; for others, as suggested by the Program Office, it may require adapting or shortening such instruments and establishing their validity and reliability. The panel also believes that the Program Office needs greater familiarity with the constructs and measures to help in guiding the discussion of the various working groups and committees to obtain concrete, pragmatic advice for implementation of the NCS. The Program Office needs to focus those discussions to obtain timely input and measurement suggestions. In contrast, the panel found that the working groups, while very thoughtful, have produced reports and documents that are relatively remote from near-term issues of implementation of the NCS.

Needed Oversight and Outside Review

In addition to the need for added depth and breadth of expertise in the NCS Program Office, the panel believes that changes are needed to strengthen the oversight and review structures for the NCS. Specifically, because of the complexity, cost, and scientific importance of the NCS and the wide range of expertise it requires in such areas as sampling, survey methods, environmental exposure measurement, health disparities and health phenotypes, the panel believes that an authoritative, multidisciplinary oversight structure for the NCS is required to ensure that the decisions of the Program Office are appropriately vetted in all relevant areas. In addition, the panel deems it critical that regular comprehensive reviews by an independent outside group, with appropriate multidisciplinary expertise, be conducted roughly every 3 years in order to ensure that the Main Study operates cost effectively to maximize the scientific utility of the information. At present, there is no known provision for such reviews.

CONCLUSION 6-1: Cost-effective and scientifically grounded operation of the National Children's Study (NCS) Main Study requires a broader

and deeper base of scientific expertise than currently exists in the NCS Program Office; an authoritative multidisciplinary oversight structure to ensure that the decisions of the Program Office are appropriately vetted in all relevant areas of expertise; and a provision for periodic comprehensive reviews of the study by an independent outside group.

APPROACHES TO CONSIDER

The panel is not the appropriate body to make specific recommendations to address the three areas of concern it has identified regarding the scientific leadership and decision making structure for the NCS. Instead, we offer below some approaches for consideration in each area—scientific expertise, oversight, and periodic outside review.

Methodological and Substantive Expertise

To develop a broader and deeper base of scientific expertise, the panel urges the NCS Program Office to consider ways and means to bring in outside scientists with relevant backgrounds and experience to supplement its in-house staff. The NCS has established the National Children's Study Scholars Program,[3] but it is limited to federal employees, and the positions are not salaried and so have to be supported by the scientist's sponsoring federal agency. To engage the expertise it needs, the NCS could provide support for appointments in the Program Office of up to several years through such mechanisms as interagency personnel agreements (IPAs) for scientists with needed expertise, not only from other federal agencies, but also from academic institutions and private research organizations.

Requiring a commitment to spend full time at the NCS for an extended period of time could be a deterrent to attracting the very best research scientists from academia. To overcome this barrier, the Program Office could also establish a competitive contract mechanism to select and fund outstanding academic research scientists to collaborate part time (e.g., 20-40 percent) on the NCS to provide expertise and possibly function as co-principal investigators for project domains. Such contracts, to be most effective, are best handled with individuals separately from other contracts for such activities as data collection.

Other approaches could also be considered to strengthen the base of scientific expertise that is available to the NCS. For example, it could be useful to increase and formalize collaborations, at the levels of both programs and individual scientists, between the Program Office and the intramural research programs of the NICHD, the National Institute of Environmental and Health

[3]For details, see http://www.nationalchildrensstudy.gov/opportunities/Pages/scholars.aspx [April 2014].

Sciences (NIEHS) and other institutes of the NIH; with CDC, including the National Center for Health Statistics (NCHS) and the National Center for Environmental Health; and with EPA.

RECOMMENDATION 6-1: The National Institute of Child Health and Human Development (NICHD) should consider and implement one or more approaches to enhance the scientific expertise of the National Children's Study (NCS) Program Office by recruiting experts in relevant fields from within the National Institute of Health, other federal agencies, and outside government. In addition, NICHD should consider contracting with experts outside of government to work part time on the NCS as a means to bolster the scientific expertise that is focused on the NCS.

Authoritative Oversight

To improve the oversight structure for the NCS and to ensure that the study design and implementation protocols are as scientifically based and cost-effective as possible, the panel suggests a model for study oversight, which builds on the language in the Children's Health Act of 2000.

Under this model, the NIH would establish a scientific management group,[4] the members of which would have the authority to review and approve the study hypotheses, design, methods, measures, cost, and instruments on an ongoing basis. Such an interagency scientific management group would include senior-level scientists from the NICHD, NIEHS, EPA, and CDC, selected by their agencies, who would be assigned to contribute substantial effort in a review capacity for the NCS, at least until the Main Study is fully launched. The management group could also consider selecting some nongovernmental experts to serve on the committee. As the Main Study progresses, the composition of the committee and level of effort could change in light of different needs for expertise to address different developmental milestones in children's health and development.

This group would make decisions about the overall design and operation of the NCS. For greater effectiveness, responsibility for specific aspects could be assigned to the participants from particular agencies within an overall agreed-to framework: for example, sample design and overseeing field operations to NCHS, environmental exposure measures to NIEHS and EPA, and child development to NICHD. This structure would be responsive to Section 1004b of the legislation and would enhance scientific legitimacy of the NCS.

[4]A "scientific management group" is equivalent to the concept of an independent oversight committee that is included in some of the recommendations earlier in this report. An interagency scientific management group is one example of how an independent oversight committee could be established.

This administrative structure could be implemented by changing the current Interagency Coordinating Committee from being an advisory, coordinating body to a study oversight committee with approval authority subject to final review by the NIH director. Alternatively, a scientific management group could be established in addition to the Interagency Coordinating Committee with higher level agency leaders who would focus on the oversight function. Such a committee should have unrestricted access to NCS processes and data that are the basis for study design decisions by the NCS Program Office.

Either of the approaches to establishing an appropriate oversight structure just outlined would be responsive to the Children's Health Act of 2000. In addition, the NCS Federal Advisory Committee would continue in operation, advising the scientific management group and expecting substantive responses to its recommendations.

RECOMMENDATION 6-2: The National Institutes of Health should strengthen the oversight and leadership of the National Children's Study (NCS) by establishing an oversight scientific management structure to include a full range of relevant expertise, with review and approval authority for NCS design and major management decisions.

Periodic Outside Review

For such a complex, long-running, and costly enterprise as the NCS Main Study, it is important to have periodic formal outside scientific review in addition to strengthening the scientific expertise devoted to the study, providing for authoritative ongoing oversight, and continuing the standing federal advisory committee. There are several possible mechanisms for accomplishing such reviews, which should generally be conducted every few years and involve not only review of documents, but also meetings with program and contractor staff, field visits to observe data collection and processing, and other activities as necessary to ensure a comprehensive review.

Models for such outside review include the use of a specially appointed visiting committee, such as the National Science Foundation convenes every few years for reviews of each of its major programs, or the appointment of a study section as used by NIH, or the commissioning of a qualified independent organization to conduct the review. The review group, however constituted, would report to the oversight scientific management group and the NIH director and its reports should be publicly released.

RECOMMENDATION 6-3: The National Children's Study (NCS) Program Office should establish a mechanism, such as a study section like those in the National Institutes of Health or qualified independent organization, to conduct periodic comprehensive outside scientific reviews

of the design and operations of the Main Study. To facilitate the work of such a committee and transparency for the study more generally, the NCS Program Office should promptly post on its Website all scientific studies conducted for the NCS.

LESSONS FROM OTHER STUDIES

The panel notes that large-scale, long-running longitudinal studies in the United States are run either by statistical agencies with extensive survey and statistical experience (e.g., the National Longitudinal Surveys of the Bureau of Labor Statistics, which contracts for the data collection) or by university-based centers that have one or more principal investigators who bring deep scientific and survey design knowledge to the project with funding from a federal grant or cooperative agreement (e.g., the Health and Retirement Study and the Panel Study of Income Dynamics at the University of Michigan and the National Longitudinal Study of Adolescent Health at the University of North Carolina). We also note that three large British birth cohort studies are housed in the same university-based research institute. We know of no example in which a funding agency, such as NICHD, not only provides the funding for a survey but also controls the design decisions and overall operations with contracts for data collection and other functions. The NCS Program Office obtains advice from researchers and relevant federal agencies, but that advice is not binding for design or operational decisions.

In a review of six longitudinal birth cohort studies, Golding (2009) commented on management structures.[5] Four of those reviewed were housed in a university, and two were managed by government agencies. Golding (2009, p. 27) suggests that "basing the director and the study in a university is likely to be the best way to ensure the study's scientific and ethical integrity." However, other national birth cohort studies not mentioned by Golding also are managed by government agencies. The panel did not reach a consensus on whether to recommend one model or the other, but the majority of panel members generally believe that both management models could be effective. The panel judges that the current management structure is not functioning effectively and encourages the NIH director to carefully reconsider all options before settling on a structure that is most likely to result in a more scientifically based and cost-effective NCS main study.

[5]One of the studies reviewed was the NCS; two other studies, the Danish National Birth Cohort and the Norwegian Mother and Child Cohort Study, enrolled 100,000 births by 2002 and 2008, respectively.

7

Summary, Conclusions, and Recommendations

This chapter brings together the panel's conclusions and recommendations from the preceding chapters for the design of the National Children's Study (NCS) Main Study. We present them in the context of the overall conceptual framework proposed for the NCS: to serve as a "data platform," with a focus on child health and development and the study design principles that flow from that framework. In the first section below, we present that framework and our findings on how the current design of the NCS meets the principles of the framework.

The chapter then summarizes the panel's assessment of the nine issues it was specifically asked to address regarding the design for the NCS:

1. the national probability sample's overall sample size and design;
2. the size of the supplemental convenience sample(s);
3. the optimal use of sibling births;
4. the relative size of the prenatal and birth strata in the probability sample;
5. the use of hospitals and birthing centers as the primary sampling units;
6. the use of health care providers to refer prospective participants;
7. the proposed approach to define and characterize health disparities;
8. the proposed approach to assess health and developmental phenotypes; and
9. the proposed study visit schedule, with emphasis on more frequent data collection in pregnancy and early childhood.

These nine issues are discussed under two headings—the design, size, and composition of the Main Study probability sample (issues 1-6) and the content and visit schedule (issues 7-9). The final two sections of this chapter cover the costs of data collection and the overall leadership of the NCS.

It is difficult to judge the impact of the panel's recommendations on the cost or timing of the Main Study. It may be that the recommendations from Chapters 2 through 5 will not cost more than the opaque process currently underway, especially given the commitment to pilot testing of the NCS Program Office. However, the panel received no hard evidence that there is a specific plan driving the work to launch the Main Study and hence no way to determine the impact of the panel's recommendations on cost and timing. The recommendations to enhance the scientific expertise, oversight, and periodic outside review of the NCS may result in delays in implementation of the Main Study. However, the panel believes that the quality, utility and cost-effectiveness of the Main Study will ultimately be enhanced thereby.

CONCEPTUAL FRAMEWORK AND IMPLICATIONS FOR DESIGN

The panel endorses the overall conceptual framework proposed for the NCS, in which it is to function as a data collection platform with a focus on health and development (see Chapter 2). From this conception of the NCS as a data platform, the panel identified the following overarching principles that are important to reflect in the study design. It then considered the proposed NCS design in light of the principles.

- **Principle**: A scientific framework that encompasses current and anticipates future domains of high-priority scientific inquiry is needed to guide key study design elements, such as the target population, the sampling strategy, and the schedule and content of data collection.
    - **Finding:** The framework for the NCS is not currently as well developed as required to meet the principle. In addition, the specific design feature to recruit equal numbers of cases into the sample prenatally and at birth does not fully reflect the growing consensus in the scientific literature on the importance of prenatal influences on child health and development.
- **Principle:** Scientifically robust exemplar hypotheses are needed to guide sample design and early-wave data collection, while decisions about data to be collected in later waves should leave room to take into account new hypotheses that emerge over the course of the study.
    - **Finding:** The proposed exemplar hypotheses for the NCS are not currently as well developed as required to guide sample design and data collection in the early waves, nor is there a long-range plan spelling out lines of inquiry that must be pursued early on to

support development of new instruments for collection of information later in the project.

- **Principle:** A probability sample ensures that results generalize to the population from which the sample is drawn.
  - **Finding:** The proposed design largely incorporates this principle, with the exception of its convenience samples.
- **Principle:** A stratified national sample in which children have an approximately equal chance of selection is required to support multiple goals. For the NCS, these goals include estimating relationships between exposures and health outcomes, analyzing health disparities, and attaining representation of children in key demographic and geographic subgroups roughly in proportion to their representation in the population.
  - **Finding:** Although the proposed NCS design largely incorporates the study's goals, its stratification plan needs to be carefully evaluated to ensure that it responds to research on the effects of exposures on health outcomes, as well as health disparities research priorities, as mandated in the Children's Health Act of 2000.
  - **Finding:** The panel was not provided with sufficient information to determine if further reductions in sample size (beyond those in the current design) that may become necessary due to costs would jeopardize the study's analytical requirements.
- **Principle:** As large a sample size as possible within budget constraints is needed to provide statistical power for current and future scientific discoveries.
  - **Finding:** The proposed NCS design largely reflects this principle, with a proposed national probability sample of 90,000.
  - **Finding:** The proposed supplemental samples of 10,000 do not add sufficient value to the study to warrant their inclusion, and they detract from the national probability sample's potential size and prenatal coverage.
- **Principle:** Scientific quality is enhanced by using the most valid and standardized data collection measures and methods that are feasible while maintaining sufficient flexibility to assess emerging domains of scientific inquiry.
  - **Finding:** The panel was not provided sufficient information with which to evaluate this aspect of the NCS design.
- **Principle:** Scientific discovery is enhanced when the potential for future innovations in measurement is incorporated into the study.
  - **Finding:** The plan for collection and storage of biological and environmental samples meets this principle and is appropriate to make them available for future investigations.

  - **Finding:** Details about how potential innovations will be adopted were not provided to the panel.
- **Principle:** Discoveries related to health conditions are facilitated by a dynamic conception of health and disease, which calls for measuring health status, disease conditions, symptoms, and behaviors rather than just existing disease categories.
  - **Finding:** The proposed design embraces this principle, but it does not provide sufficient details for the panel to assess whether the burden imposed on respondents by the additional questions on the conditions and symptoms is excessive.
  - **Finding:** The process for specifying the measures to be collected appears to be large, unwieldy, and unsuitable for field implementation, and the operationalization of this process would benefit from timely and transparent scientific consultation.
- **Principle:** Discovery is facilitated if data are released as early and as completely as possible, with due regard for the protection of confidentiality.
  - **Finding:** The proposed design endorses this principle but would profit from lessons provided by data release schedules and methods followed in other national studies to achieve timely release.
- **Principle:** Transdisciplinary discovery and statistical sophistication are enhanced when all relevant scientific expertise is integrated into the project management structure.
  - **Finding:** The NCS design, as described, indicates a lack of sufficient scientific expertise, which is a major weakness of the study.
- **Principle:** The study design should be as cost effective and efficient as possible.
  - **Finding:** The panel was not provided sufficient information with which to evaluate this aspect of the NCS design.

The above conceptual framework, resulting design principles, and panel findings, as well as the scientific literature on children's health and development, lead to the panel's recommendations regarding the overall design of the NCS.

RECOMMENDATION 2-1: The scientific framework for the National Children's Study should be based on current understanding of the determinants of children's health and development and an informed consideration of the likely future trajectory of scientific discovery. The paradigms of developmental biology and life-course epidemiology, coupled with findings from other social and behavioral science research on the prenatal and early life periods, should guide development of the design for the Main Study.

RECOMMENDATION 2-2: In order to facilitate scientific discovery during and after National Children's Study data are gathered, the Main Study should use a national probability sample with the largest feasible sample size and an approximately equal probability of selection design, and it should recruit nearly all of the cohort as early in pregnancy as possible.

RECOMMENDATION 2-3: In order to facilitate scientific discovery during and after National Children's Study (NCS) data are gathered, the Main Study should use valid and standardized data collection measures and methods, while maintaining flexibility to revise or develop new instruments. The NCS should also use state-of-the-art procedures to collect, archive, and provide access to biological and environmental specimens for future analyses.

RECOMMENDATION 2-4: The proposed strategy for the National Children's Study Main Study to collect detailed data on children's health status, conditions, symptoms, and behaviors should be followed to the extent possible, taking into account constraints of costs, operational feasibility, and the need to not overburden respondents.

The panel further offers two recommendations about the proposed supplemental samples for the NCS.

RECOMMENDATION 2-5: While the panel appreciates the possible scientific value of gathering preconception exposure information on 5,000 first-birth children as part of the National Children's Study Main Study, this supplemental sample should be dropped because of high costs, the lack of any evidence of the value of such a sample, the lack of detailed plans for both selection and analysis, and potential limitations in the proposed data collection schedule.

In making this recommendation, the panel also took into consideration the loss of the opportunity to recruit more prenatal cases if the preconception group is retained.

RECOMMENDATION 2-6: The other supplemental convenience samples proposed for the National Children's Study Main Study should be dropped from the design, including samples of children exposed to natural disasters or geographically defined environmental exposures, samples of additional members of disadvantaged groups, and samples of siblings born outside the 4-year birth window. The potential added value of the supplemental sample cases is less than the value of the additional cases

in the probability sample they would replace, specifically, the value of the additional prenatal cases in the probability sample.

Regarding the scientific consensus on the importance of beginning data collection during the prenatal period cited in Recommendation 2-1 (above), the panel notes that the Program Office did not provide a scientific rationale to support the proposed change to enroll one-half of the probability sample at birth instead of enrolling as many cases as possible prenatally, as in the original design. The Program Office suggested that resource constraints led to this design, but it did not provide cost estimates for its proposed design or for any alternative design models. The panel conducted its own cost analysis for recruitment and data collection under alternative designs (see Chapter 5 and Appendix B): it demonstrates that a 100 percent national probability sample, with as much prenatal recruitment as possible,[1] has approximately the same field cost and yields nearly as many total cases as the proposed sample of 45 percent recruitment prenatally, 45 percent recruitment at birth, and 10 percent convenience or supplemental cases.

PROBABILITY SAMPLE DESIGN, SIZE, AND COMPOSITION

In the proposed design for the probability sample as described in materials provided to the panel, the first design decision is to define the study's target population: it is to consist of all births in the United States (50 states and the District of Columbia) but excluding births to women who do not deliver in hospitals or birthing centers and births to women at hospitals with too few births to be included on the sampling frame. A second design decision involves the overall size of the probability sample.

CONCLUSION 3-1: The panel endorses the proposed target population of all births in the United States during a specified time period consisting of 4 full calendar years, as well as the proposed sample exclusions from this target population.

CONCLUSION 3-2: Because of the lack of explicit hypotheses in the study design, it is not possible for the panel to judge whether the proposed sample size is justified on the basis of the study's objectives.

A large appropriately stratified national probability sample in which children have an approximately equal chance of selection would be one that

[1]As noted in Chapter 4, birth recruitment would be needed for the relatively small number of women who do not receive prenatal care or who receive prenatal care from a provider not on the provider list frame.

largely ensures that children in key demographic and geographic subgroups are represented roughly in proportion to their representation in the population. Such a sample could achieve the needed geographic dispersion within key demographic groups to facilitate analysis of health disparities. However, while stratification is a key mechanism for assuring this chance of selection, little information was available concerning what kind of stratification will be possible. In addition, other studies have found significant differential attrition among subgroups of particular interest to the assessment of health disparities: this is a key concern given the 21-year life of the NCS. Differential attrition will affect the ultimate composition of the sample and may reduce its value for assessing health disparities if not anticipated and addressed using oversampling.

CONCLUSION 3-3: By adopting an equal probability of selection design for the National Children's Study, it is likely that the sample sizes for a number of subgroups of interest may be inadequate for some important types of analysis. These subgroups are likely to include minorities in the U.S. population who are known to be on the negative side of health disparities and to have higher attrition in longitudinal studies. However, the absence of explicit study hypotheses and objectives makes it difficult to identify these important population subgroups and their associated sample size requirements.

RECOMMENDATION 3-1: The National Children's Study Main Study sample should be stratified by characteristics that will achieve variability in socioeconomic status within important population groups to support analysis of health disparities, as well as achieving variability in environmental exposures and geography to support analysis of relationships between exposures and health outcomes.

Insufficient information was provided to the panel to assess the coverage, feasibility, and other aspects of the first stage of the proposed design—using a list of hospitals as primary sampling units (PSUs)—because development of the sampling plan had not been completed as of February 2014. With the proposed design, the secondary sampling stage would be prenatal care providers from the sampled hospitals, split into two strata: the prenatal stratum and the birth stratum. The ultimate sampling stage would be the sampling of pregnant women from selected providers or the sampling of women just after birth at selected hospitals.

CONCLUSION 3-4: The panel has not been provided with sufficient detail on the planned hospital-based sample design and recruitment strategy to judge their merits and scientific validity or determine potential coverage bias and the availability of appropriate stratification variables.

CONCLUSION 3-5: The panel has not been provided with sufficient justification for moving to hospital-based primary sampling units from the sampling approach previously proposed by the National Children's Study (NCS) for discussion at the 2013 NCS Workshop (see National Research Council and Institute of Medicine, 2013) and based on Vanguard Study pilot testing—namely, county-based primary sampling units with hospitals as secondary sampling units and providers as third-stage sampling units.

Assessment of the proposed sample design, when completed, should include comparisons with the previous design. Any comparison needs to include a cost-effectiveness analysis of the options and an assessment of the ability to ensure coverage and to control for such characteristics as race and ethnicity, socioeconomic status, age, and marital status to ensure the sample will support evaluation of health disparities.

Because a geographic-based first stage sample design has already been developed and would need only to be updated for population changes and because it appears feasible to sample prenatal care providers within geographic areas based on the Vanguard Study experience, the panel questions the decision to move to a hospital-based approach.

In its Vanguard Study, the NCS tested both provider-based recruitment and provider-based sampling approaches. While the panel was provided with limited information about what was done and how well it worked, the available information does indicate that provider-based sampling followed by recruitment of women is feasible. However, the panel is concerned about the possibility of high rates of nonparticipation, particularly by women in groups important for understanding health disparities: it will be important for the NCS to monitor participation rates by important categories, such as race and ethnicity, socioeconomic status, age, and marital status.

CONCLUSION 3-6: Assuming that participation in the National Children's Study Main Study follows the patterns in the Vanguard Study, the cumulative response rate to birth for the prenatal stratum would be between 28 and 32 percent, and the rate to age 12 would be 22 to 25 percent—very small fractions of the eligible sample. The cumulative response rate to age 21 would be 18 to 21 percent. A thorough analysis of nonresponse bias is clearly indicated, and in any case will be required by the U.S. Office of Management and Budget.

CONCLUSION 3-7: The high rates of cumulative nonresponse expected in the National Children's Study pose a severe risk for nonresponse bias that may not be mitigated by weighting adjustments, potentially making some study results invalid.

A range of other aspects of the current sampling plan were considered by the panel.

CONCLUSION 3-8: Enrolling siblings as members of the National Children's Study sample provides many analytic advantages, most prominently the gathering of preconception exposure information for second- and higher-order births. The panel endorses current plans to recruit siblings born after the initially recruited child—but only within the 4-year recruitment interval associated with the original primary sampling unit for the target birth—and to continue to follow these children until age 21.

CONCLUSION 3-9: Weight adjustment and screening are viable options for accounting for the fact that subsequent siblings have more than one way to enter the sample. The panel was not provided sufficient information to recommend one over the other. In either case, detailed information on prior births to the mother will need to be collected.

CONCLUSION 3-10: As of February 2014, the currently proposed hospital-based sample design for the National Children's Study had not been sufficiently developed or documented to support an evaluation.

CONCLUSION 3-11: The identification, sampling, and recruitment of women at the time of birth has not been sufficiently pilot tested, using a representative set of hospitals, to support any conclusion about this feature of the design.

RECOMMENDATION 3-2: A detailed plan for sampling, recruitment, and minimizing attrition bias for the National Children's Study (NCS) Main Study should be fully developed and evaluated by sampling and survey experts independent from the NCS and approved by the proposed independent oversight committee before the study moves forward.

CONTENT AND VISIT SCHEDULE

The proposed approach to assess health and developmental phenotypes, the proposed study visit schedule, and the development of content for the NCS reflect a conceptualization of health and development that is a substantial advance from the one that was reviewed in the previous evaluation by the National Research Council and Institute of Medicine (2008). The breadth of the conceptualization encompasses most of the issues affecting child health and development and provides many dimensions that could be linked to environmental exposures. In addition, the flexibility to use data to generate a variety of phenotypes, rather than focus on specific diagnoses, seems promising. How-

ever, needed details on the operationalization and effectiveness of these new approaches were not provided to the panel.

Exemplar hypotheses are a valuable way to guide sample design and early-wave data collection, while the data collected in later waves need to be able to adapt to hypotheses that emerge over the course of the study. Such hypotheses need to be carefully formulated in the context of the overall goals of a study.

CONCLUSION 4-1: A strategy of using a few exemplar hypotheses rather than stating a large list of hypotheses requires that the planners of the National Children's Study ensure that the exemplar hypotheses are important and scientifically robust to guide the study design and data collection.

RECOMMENDATION 4-1: Prior to proceeding with the Main Study, the National Children's Study (NCS) should develop scientifically well-grounded exemplar hypotheses that should be used to guide and evaluate decisions regarding the NCS design and data collection schedule and domains.

RECOMMENDATION 4-2: Because hypotheses will change over time, the National Children's Study should implement a strong and public process to revise and develop new exemplar hypotheses to guide future study implementation, engaging with the extramural and intramural research communities.

In addition to exemplar hypothesis, other aspects of the NCS content have not been sufficiently detailed to be used for design decisions.

CONCLUSION 4-2: While using a dynamic health phenotype concept to plan the content of the National Children's Study appears to be a promising strategy, the panel lacked sufficient information to judge whether the implementation of such an approach would be feasible given constraints imposed by respondent burden and overall study costs.

CONCLUSION 4-3: Many of the principles and concepts guiding development of the study design and the concept of having processes for developing future hypotheses and study content are consistent with the study platform framework for the National Children's Study. However, it is not clear whether and how those principles and concepts can be effectively used to design the study content.

The schedule of data collection is a key design element for any study, and it is especially important for a large-scale longitudinal study such as the NCS. Other critical elements in any study design are enrollment and the protocols for

data collection. In addition, two key elements for the NCS are exposure to a wide range of hormonal, chemical, and other environmental factors and a focus on health disparities. In all these areas, the information provided to the panel lacked sufficient information or scientific justification for the current design for the NCS Main Study. More broadly, the overall processes for developing the details needed for careful design is unclear.

CONCLUSION 4-4: The panel agrees that more intensive data collection in the early years of the National Children's Study is important, but the panel did not receive sufficient scientific justification to assess the merits of the specific data collection schedule.

RECOMMENDATION 4-3: The National Children's Study Main Study should collect data during the prenatal period at multiple times for as many of the study participants as the budget will allow.

CONCLUSION 4-5: The strategy of the National Children's Study (NCS) to enroll a substantial proportion of participants at the time of the child's birth poses substantial logistical and operational challenges that have not been adequately tested in the NCS Vanguard Study.

RECOMMENDATION 4-4: Although the panel does not endorse the current proposal for a substantial birth enrollment stratum, if the National Children's Study (NCS) Main Study retains such a stratum, the NCS should conduct a full pilot test of recruitment and data collection during the birth visit before implementation.

CONCLUSION 4-6: Exposure assessment, including collection of biological and environmental specimens during multiple study visits beginning during the prenatal period, is a critical component of the National Children's Study in addressing the mandate of the Children's Health Act of 2000 and fulfilling the study's goal to serve as a platform for future scientific inquiry.

CONCLUSION 4-7: The processes for developing content for the National Children's Study are complicated, and insufficient documentation has been provided to demonstrate that the processes will be effective.

RECOMMENDATION 4-5: The National Children's Study Program Office should document and provide justification for development of the data collection schedule, content, and methods now and going forward. The documentation should be sufficient to guide use of the study data by future researchers.

RECOMMENDATION 4-6: The National Children's Study Program Office should finalize the study visit data collection protocols that it intends to use for the Main Study (including questionnaires and other measurements), at least through age 1, and then pilot test the protocols before implementing the Main Study. The protocols and findings of the pilot tests should be peer reviewed and approved by the proposed independent oversight committee prior to initiating the Main Study.

CONCLUSION 4-8: Based on the information provided, the panel concludes that the National Children's Study plan has paid insufficient attention to how health disparities should be taken into account in the development of the schedule visit and content of the Main Study.

RECOMMENDATION 4-7: The relevance to health disparities should be an explicit criterion for selecting the constructs that will be assessed as part of the National Children's Study (NCS) Main Study, the measures that will be used to assess them, and the timing of the assessments. The NCS should obtain input from experts on health disparities in childhood as part of the documented process through which the measures for inclusion are selected and the measures should be approved by the proposed oversight committee.

A final data issue for any longitudinal study concerns the release of its data.

CONCLUSION 4-9: The panel endorses the general structure of the data distribution plans for the National Children's Study (NCS), although it fails to understand the need for a 2-year lag between the availability of analytic data and their release to the research community. Subject to confidentiality concerns, timely and complete data access are vital to maximize the scientific value of the NCS and have been achieved by other federal government surveys, which ought to serve as models for the NCS.

RECOMMENDATION 4-8: The National Children's Study should consider producing an "early release" version of the data from the Main Study that includes data collected in the early years of each wave's data collection cycle and makes those data available to analysts under the terms of restricted access data centers.

FIELD COSTS

For virtually all large-scale studies there are tradeoffs between everything one would like to do and what the budget can support. Those tradeoffs need to be made on the basis of realistic assumptions and careful analysis of the costs

of various aspects of the study. Because the NCS has not undertaken such an analysis and in order to fulfill the panel's charge for a comprehensive review of the plans for the Main Study, the panel undertook an analysis of the NCS's likely field costs. Although field costs are only part of overall study costs, they are the ones most likely to be affected by the sample design features considered by the panel.

CONCLUSION 5-1: The panel estimates that the field costs of the current plan for the National Children's Study over its first 7 years will total $1.495 billion. This estimate does not include a number of other sizable contributors to overall study costs.

CONCLUSION 5-2: The major drivers of field costs for the National Children's Study Main Study are sample size and number of interviews, with an in-person interview costing more than twice as much as a telephone interview. In contrast with contacting and gaining the cooperation of respondents, modest changes in interview length contribute minimally to cost.

CONCLUSION 5-3: For the same field costs and with the elimination of the 10,000 supplemental samples—5,000 nulliparous women and a 5,000 convenience sample—the National Children's Study could afford to enroll a predominantly prenatal probability sample of 96,000 cases with no other changes to the proposed data collection protocol.

RECOMMENDATION 5-1: Given the goal for the National Children's Study (NCS) to understand the links of environmental exposures to child health and development and its cost structure, if major reductions in the cost of the study need to be made, they should be reductions in sample size rather than exposure domains. Along with such a decision to reduce the sample size, the NCS should reconsider whether to oversample minorities in order to maintain the ability to evaluate health disparities with a reduced sample.

LEADERSHIP OF THE NATIONAL CHILDREN'S STUDY

Throughout the panel's deliberations, the NCS Program Office staff were cooperative and responsive, providing timely responses to the panel's many questions. Yet the panel repeatedly found that the answers to its questions were less than what would be needed to carry out its charge for a comprehensive review. The lack of detailed information to many basic questions about the design of the NCS and the lack of scientific justification for many of the design

decisions for the Main Study led the panel to deep concern about the overall leadership and management of the NCS.

CONCLUSION 6-1: Cost-effective and scientifically grounded operation of the National Children's Study (NCS) Main Study requires a broader and deeper base of scientific expertise than currently exists in the NCS Program Office; an authoritative multidisciplinary oversight structure to ensure that the decisions of the Program Office are appropriately vetted in all relevant areas of expertise; and a provision for periodic comprehensive reviews of the study by an independent outside group.

RECOMMENDATION 6-1: The National Institute of Child Health and Human Development (NICHD) should consider and implement one or more means to enhance the scientific expertise of the National Children's Study (NCS) Program Office by recruiting experts in relevant fields from within the National Institutes of Health, other federal agencies, and outside government. In addition, NICHD should consider contracting with experts outside of government to work part time on the NCS as a means to bolster the scientific expertise that is focused on the NCS.

RECOMMENDATION 6-2: The National Institutes of Health should strengthen the oversight and leadership of the National Children's Study (NCS) by establishing an oversight scientific management structure to include a full range of relevant expertise, with review and approval authority for NCS design decisions.

RECOMMENDATION 6-3: The National Children's Study (NCS) Program Office should establish a mechanism, such as a study section like those in the National Institutes of Health, or use a qualified independent organization to conduct periodic comprehensive outside scientific reviews of the design and operations of the NCS Main Study. To facilitate the work of such a committee and transparency for the study more generally, the NCS Program Office should promptly post on its Website all scientific studies conducted for the NCS.

References: General

Akinbami, L.J., J.E. Moorman, P.L. Garbe, and E.J. Sondik. (2009). Status of childhood asthma in the United States, 1980-2007. *Pediatrics* 123(Suppl. 3):S131-S145.

Baker, D., C. Park, C. Sweeney, L. McCormack, M. Durkin, R. Brenner, D. Dabelea, and B. Entwisle. (2014). Recruitment of women in the National Children's Study initial Vanguard Study. *American Journal of Epidemiology.* Available: http://aje.oxfordjournals.org/content/early/2014/05/01/aje.kwu062.short?rss=1 [May 2014].

Barker, D.J., P.D. Winter, C. Osmond, B. Margetts, and S.J. Simmonds. (1989a).Weight in infancy and death from ischaemic heard disease. *Lancet* 2:577-580.

Barker, D.J., C. Osmond, J. Golding, D. Kuh, and M.E. Wadsworth. (1989b). Growth in utero, blood pressure in childhood and adult life, and mortality from cardiovascular disease. *British Medical Journal* 298:564-567.

Belanger, K., S. Buka, D.C. Cherry, D.J. Dudley, M.R. Elliott, D.E. Hale, I. Hertz-Picciotto, J.L. Illuzzi, N. Paneth, J.M. Robbins, E.W. Triche, and M.B. Bracken. (2013). Implementing provider-based sampling for the National Children's Study: Opportunities and challenges. *Paediatric and Perinatal Epidemiology* 27:20-26.

Braveman, P.A., S. Kumanyika, J. Fielding, T. Laveist, L.N. Borrell, R. Manderscheid, R., and A. Troutman. (2011). Health disparities and health equity: The issue is justice. *American Journal of Public Health* 101(Suppl. 1):S149-S155.

Centers for Disease Control and Prevention. (2011). *Natality public use file*. Available: http://www.cdc.gov/nchs/data_access/vitalstatsonline.htm [May 2014].

Charles, M.A., H. Leridon, P. Dargent, B. Geay, and the Elfe team. (2011). Tracking the lives of 20,000 children: Launch of the Elfe child cohort study. *Population and Societies* 475. Available: http://www.ined.fr/fichier/t_publication/1530/publi_pdf2_pesa475.pdf [April 2014].

Cohen Hubal, E.A., T. de Wet, L. Du Toit, M.P. Firestone, M., Ruchirawat, J. van Engelen, and C. Vickers. (2013). Identifying important life stages for monitoring and assessing risks from exposures to environmental contaminants: Results of a World Health Organization review. *Regulatory Toxicology and Pharmacology* 69(1):113-124. Available: http://dx.doi.org/10.1016/j.yrtph.2013.09.008 [April 2014].

Duncan, G.J., J. Boisjoly, and K. Mullan Harris. (2001). Sibling, peer, neighbor, and schoolmate correlations as indicators of the importance of context for adolescent development. *Demography* 38(3):437-447.

Ernst, L.R. (1999). *The maximization and minimization of sample overlap problems: A half century of results*. International Statistical Institute, Proceedings, Invited Papers, IASS Topics, 168-182. Available: http://www.stat.fi/isi99/proceedings/arkisto/varasto/erns0184.pdf [April 2014].

Golding, J. (2009). A guide to undertaking a birth cohort study: Purposes, pitfalls, and practicalities. *Paediatric and Perinatal Epidemiology* 23(Suppl. 1):23-30.

Guttmacher, A.E., S. Hirschfeld, and F.S. Collins. (2013). The National Children's Study: A proposed plan. *New England Journal of Medicine* 369(20):1873-1875.

Hamilton, B.E., D.L. Hoyert, J.A. Martin, D.M. Strobino, and B. Guyer. (2013). Annual summary of vital statistics: 2010-2011. *Pediatrics* 131:548-558.

Hardy, R., and D. Kuh. (2009). The practicalities of undertaking family-based studies. In *Family Matters: Designing, Analyzing, and Understanding Family-Based Studies in Life Course Epidemiology*, pp. 181-191, D.A. Lawlor, and G.D. Mishra (eds.). Oxford: Oxford University Press.

Hofman, A., V. Jaddoe, J. Mackenbach, H. Moll, R. Snijders, E. Steegers, F. Verhulst, J. Witteman, and H. Buller. (2004). Growth, development, and health from early fetal life until young adulthood: The Generation R Study. *Paediatric Perinatal Epidemiology* 18(1):61-72.

Jaddoe, V.V., C. Duijn, O. Franco, A.J. van der Heijden, M.H. van Ilzendoorn, J.C. de Jongste, A. van der Lugt, J.P.Mackenbach, H.A. Moll, H. Raat, F. Rivandeneira, E. Steegers, H. Tierneier, A.G. Uitterlinden, F.C. Verhulst, and A. Hofman. (2012). The generation R study: Design and cohort update 2012. *European Journal of Epidemiology* 27:739-756.

Kerver, J., M. Elliott, G. Norman, R. Sokol, D. Keating, G. Copeland, C. Johnson, K. Cislo, K. Alcser, S. Kruger-Ndiave, B. Pennell, S. Methta, C. Joseph, N. Paneth, and the MANCS Executive Committee. (2013). Pregnancy recruitment for population research: The National Children's Study Vanguard experience in Wayne County, Michigan. *Paediatric and Perinatal Epidemiology* 27:303-311

Landrigan, P.J., and A. Miodovnik. (2011). Children's health and the environment: An overview. *Mount Sinai Journal of Medicine* 78(1):1-10.

Landrigan, P.J., L. Trasande, L.E. Thorpe, C. Gwynn P.J. Lioy, M.E. D'Alton, H.S. Lipkind, J. Swanson, P.D. Wadhwa, E.B. Clark, V.A. Rauh, F.P. Perera, and E. Susser. (2006). The National Children's Study: A 21-year prospective study of 100,000 American children. *Pediatrics* 118(5):2173-2186.

LaVeist, T.A. (2005). Disentangling race and socioeconomic status: A key to solving health disparities. *Journal of Urban Health* 82(Suppl. 3):26-34.

Lawlor, D.A., A.M. Andersen, and G.D. Batty. (2009). Birth cohort studies: Past, present, and future. *International Journal of Epidemiology* 38:897-902.

MacDorman, M.F., M.S. Matthews, and E. Declercq. (2014). *Trends in out-of-hospital births in the United States, 1990-2012.* National Center for Health Statistics. NCHS Data Brief No 14. Available: http://www.cdc.gov/nchs/data/databriefs/db144.htm [March 2014].

Magnus, P., L.M. Irgens, K. Haug, W. Nystad, R. Skjaerven, C. Stoltenberg, and the Moba Study Group. (2006). Cohort profile: The Norwegian Mother and Child Cohort Study (MoBa). *International Journal of Epidemiolology* 35(5):1146-1150.

Michael, R.T., and C.A. O'Muircheartaigh. (2008). Design priorities and disciplinary perspectives: The case of the US National Children's Study. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 171(2):465-480.

National Center for Health Statistics. (2014). *NHANES data release and access policy.* Available: http://www.cdc.gov/nchs/data/nhanes/nhanes_release_policy.pdf [February 2014].

National Center for Health Statistics (no date). *NHANES data release and access policy.* Available: http://www.cdc.gov/nchs/data/nhanes/nhanes_release_policy.pdf [February 2014].

National Research Council and Institute of Medicine. (2000). *From Neurons to Neighborhoods: The Science of Early Childhood Development.* Committee on Integrating the Science of Early Childhood Development. J.P. Schonkoff, and D.A. Phillips (eds.). Board on Children, Youth, and Families. Commission on Behavioral and Social Science and Education. Washington, DC: National Academy Press.

National Research Council and Institute of Medicine. (2004). *Children's Health, the Nation's Wealth: Assessing and Improving Child Health.* Committee on Evaluation of Children's Health: Measures of Risk, Protective, and Promotional Factors for Assessing Child Health in the Community. Washington, DC: The National Academies Press.

National Research Council and Institute of Medicine. (2008). *The National Children's Study Research Plan: A Review.* Panel to Review the National Children's Study Research Plan. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Board on Population Health and Public Health Practice, Institute of Medicine. Washington, DC: The National Academies Press.

National Research Council and Institute of Medicine. (2013). *Design of the National Children's Study: A Workshop Summary*. N. Kirkendall, *Rapporteur*. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Board on Children, Youth and Families, Institute of Medicine. Washington, DC: The National Academies Press.

NICHD. See separate reference list for National Institute of Child Health and Human Development.

Nybo Andersen, A-M,. M. Madsen, and D.A. Lawlor. (2009). Birth cohorts: A resource for life course studies. In *Family Matters: Designing, Analyzing, and Understanding Family-Based Studies in Life Course Epidemiology*, pp. 181-191, D.A. Lawlor, and G.D. Mishra (eds.). Oxford: Oxford University Press.

Olsen, J. (2012). Nine months that last a lifetime. Experience from the Danish National Birth Cohort and lessons learned. *International Journal of Hygiene and Environmental Health* 215:142-144.

Olsen, J., M. Melbye, S.F. Olsen, T.I.A Sørensen, P. Aaby, A.-M. Nybo Andersen, D. Taxbol, K.D. Hansen, M. Juhl, T.B. Schow, H. Toft Sorensen, J. Andresen, E.L. Mortensen, A.W. Olesen, and C. Sondergaard. (2001). The Danish national birth cohort: Its background, structure, and aim. *Scandinavian Journal of Public Health* 29:300-307.

Paneth, N. (2013). Restoring science to the National Children's Study. *Journal of the American Medical Association* 309(17):1775-1776.

Pearson, H. (2011). Epidemiology: Study of a lifetime. *Nature* 471:20-24.

Seith, D., and E. Isakson. (2011). *Who Are America's Poor Children? Examining Health Disparities Among Children in the United States*. National Center for Children in Poverty, Mailman School of Public Health, Columbia University. Available: http://www.nccp.org/publications/pdf/text_995.pdf [March 2014].

U.S. Environmental Protection Agency. (2010). *Workshop on Optimizing Exposure Metrics for the NCS.* Available: http://www.epa.gov/nerl/download_files/documents/national_childrens_study_june_2010.pdf [February 2014].

Vrijheid, M., M. Casas, A. Bergström, A. Carmichael, S. Cordier, M. Eggesbø, E. Eller, M.P. Fantini, M.F. Fernández, A. Fernández-Somoano, U. Gehring, R. Grazuleviciene, C. Hohmann, A.M. Karvonen, T. Keil, M. Kogevinas, G. Koppen, U. Krämer, C.E. Kuehni, P. Magnus, R. Majewska, A.M. Andersen, E. Patelarou, M.S. Petersen, F.H. Pierik, K. Polanska, D. Porta, L. Richiardi, A.C. Santos, R. Slama, R.J. Sram, C. Thijs, C. Tischer, G. Toft, T. Trnovec, S. Vandentorren, T.G. Vrijkotte, M. Wilhelm, J. Wright, and M. Nieuwenhuijsen. (2012). European birth cohorts for environmental health research. *Environmental Health Perspectives* 120(1):29-37.

Wadman, M. (2009). Children's study fights to survive. *Nature News* 462:20-21. Available: http://www.nature.com/news/2009/091103/full/462020a.html [April 2014].

Weisberg, H.F. (2005). *The Total Survey Error Approach: A Guide to the New Science of Survey Research*. Chicago: University of Chicago Press.

Williams, D.R., and M. Sternthal. (2010). Understanding racial/ethnic disparities in health: sociological contribution. *Journal of Health and Social Behavior* 51(Suppl. 1):S15-27.

Wooldridge, J.M. (2012). *Introductory Econometrics: A Modern Approach, 5th Edition.* Independence, KY: Cengage Learning.

Yang, D., G. Howard, C.S. Coffey, and J. Roseman. (2004). The confounding of race and geography: How much of the excess stroke mortality among African Americans is explained by geography? *Neuroepidemiology* 23(3):118-122

References: National Institute of Child Health and Human Development

NOTE: The documents that have been published can be identified by a Website address. The documents that were provided directly to the panel are available in the Public Access File (PAF) in the Public Access Records Office at the National Academies.

National Institute of Child Health and Human Development (2012). *National Children's Study: Potential Sampling Strategies*. Draft 3.0 April 16, 2002. Available: http://www.nationalchildrensstudy.gov/about/organization/advisorycommittee/Pages/Potential-Sampling-Strategies-NCS-4-16-12-Draft-3.0.pdf [May 2014].

National Institute of Child Health and Human Development. (2013a). *The National Children's Study Federal Advisory Committee Briefing Book*. National Children's Study Program Office. Available: http://www.nationalchildrensstudy.gov/about/organization/advisorycommittee/Pages/july_2013.aspx [December 2013].

National Institute of Child Health and Human Development. (2013b). *The National Children's Study: Description of Proposed Methods for the Main Study.* National Children's Study Program Office. Available: http://www.nationalchildrensstudy.gov/research/workshops/Pages/IOM-Workshop-on-the-Design-of-the-National-Childrens-Study-and-Implications-for-the-Generalizability-of-Results.aspx [December 2013].

National Institute of Child Health and Human Development. (2013c). *The National Children's Study Federal Advisory Committee Briefing Book* (October 22). National Children's Study Program Office. Available: http://www.nationalchildrensstudy.gov/about/organization/advisorycommittee/Pages/october_2013.aspx [December 2013].

National Institute of Child Health and Human Development. (2013d). *IOM/NRC Review: Design of the National Children's Study and Implications for Generalizability of Results*. Responses to questions from the review committee (October 25). National Children's Study Program Office. Available: PAF.

National Institute of Child Health and Human Development. (2013e). *Miscellaneous Questions for NCS from Study Panel* (October 23). National Children's Study Program Office. Available: PAF.

National Institute of Child Health and Human Development. (2013f). *Responses to Questions on Design of the NCS Main Study* (November 8). National Children's Study Program Office. Available: PAF.

National Institute of Child Health and Human Development. (2013g). *NCS Responses to Panel Questions* (November 26). National Children's Study Program Office. Available: PAF.

National Institute of Child Health and Human Development. (2013h). *Questions and Responses to NCS IOM Review Panel Questions 11-27-2013(as revised on December 17, 2013).* National Children's Study Program Office. Available: PAF.

National Institute of Child Health and Human Development. (2013i). *PSU Study* (December 17). National Children's Study Program Office. Available: PAF.

National Institute of Child Health and Human Development. (2013j). *Cost questions* (December 4). National Children's Study Program Office. Available: PAF.

National Institute of Child Health and Human Development. (2013k). *The National Children's Study Data Access and Confidentiality—Concepts of Operations v1.0.* National Children's Study Program Office. Available: http://www.nationalchildrensstudy.gov/about/overview/Pages/NCS-CONOPS-Data-Access-and-Confidentiality.pdf [February 2014].

National Institute of Child Health and Human Development. (2014a). *Cost Assessment Responses* (January 15). National Children's Study Program Office. Available: PAF.

National Institute of Child Health and Human Development. (2014b). *National Children's Study Federal Advisory Committee Briefing Document*. National Children's Study Program Office. Available: http://www.nationalchildrensstudy.gov/about/organization/advisorycommittee/Pages/January-2014-NCSAC-Briefing-Document.pdf [February 2014].

Appendix A

Communications Between the Panel and the NCS Program Office

During the course of its work, the National Children's Study (NCS) Program Office sent several background documents to the panel, and the panel sent to the Program Office several sets of questions about the current proposed design of the NCS. All questions and responses were sent and received by email between the panel's study director and the staff person designated as the point of contact or the director of the Program Office. This appendix details the timing of those panel communications with the Program Office. Table A-1, which provides a summary of the information flows, also lists information requested by the panel that was not provided by the Program Office.

August 2013 The first set of questions was sent to the Program Office on August 1, before the first panel meeting. These questions focused on the sample design. The primary purpose of this set of questions was to facilitate discussion during the panel's first meeting and to inform the Program Office that the panel was interested in seeing detailed descriptions of proposed methodologies and data used in justifying decisions. On August 8, the Program Office provided the panel with its briefing document (NICHD, 2013b). On August 16, during the part of the panel's meeting that was open to the public, the panel questioned the cost analysis used by NCS to justify its decision about having one-half of the probability sample recruited at birth and also asked when we would receive details about the hospital frame and sample.

September 2013 After the initial public meeting with Program Office staff, the panel prepared 11 sets of questions, 9 of which focused on the 9 topics

in the charge to the panel. The other two sets of questions concerned issues the panel deemed important: NCS plans to disseminate data and the NCS's planned use of interviewing modules that subsets of respondents would receive. All 11 sets of questions were sent to the Program Office by September 9. On September 30, the panel recognized gaps in the questions that had been sent on September 9 and sent another set of questions—11 in all—to NCS by email. On September 30, just prior to the government shutdown and at the urging of the panel, the Program Office provided draft responses to the 11 question sets sent to the Program Office on September 9. Given the volume and detailed nature of the questions, the Program Office stated that its responses should be considered preliminary, were provided so that the panel could continue its work during the shutdown, and might be revised in the future.

October 2013 The NCS Program Office delivered a preliminary version of responses to the questions sent on September 30, and a final version of its responses to the panel's original 11 sets of questions on October 23 (NICHD, 2013d). This document also included summary information about NCS plans. On October 30 the panel sent its fourth set of questions to NCS. These questions focused on information not previously provided by NCS, in addition to asking for clarification and detail on a number of points made in the responses that the panel had already received.

November 2013 On November 9, the panel received final responses from NCS (NICHD, 2013e) to the second set of 11 questions, which had been sent to NCS on September 30. Responses to most of the October 30 questions were received in three parts on November 8 (NICHD, 2013f), November 26 (NICHD, 2013g), and November 27 (a draft of responses to most of the remaining questions.) In a number of instances responses were either still not complete, or indicated a lack of understanding of the question. For example, in response to "the panel would like to see details about how differing four year recruitment periods will be reconciled during the estimation process," NCS described the operational need to roll out the sample over time.

On November 19 the panel's cost consultants met with NCS leadership to discuss assumptions for the cost analysis. On November 23 the panel's study director reminded NCS of the panel's need for cost information, stating that cost data are critical to the assessment of the NCS proposed design of the Main Study because cost is cited by NCS as the rationale for key decisions.

In response to cost questions, NCS stated that Program Office staff had met with the panel's cost consultants and had given them all the required information.

December 2013 The panel's cost consultants did not agree with the Program Office's assessment and formulated a set of cost questions that were communicated to the Program Office on December 6. On December 17, the panel received responses to some but not all of those cost questions. After pointing out that the responses were incomplete, a more complete set of responses was received on December 20 (NICHD, 2013j). However, even in this document, for example, the question "What are the cost metrics for these pilot efforts?" (p. 2) was not answered by the Program Office.

On December 17, NCS provided the panel with a summary of the preliminary report concerning the design for hospitals as primary sampling units (NICHD, 2013i) but not the report itself. NCS also provided the final version of responses to the last set of questions, which had been sent on October 30 (NICHD, 2013h).

January 2014 At the invitation of the panel, NCS provided comments on the panel's cost analysis (NICHD, 2014a).

Other Information The Program Office provided the panel with background documents prepared for the three meetings of the NCS Advisory Committee that were held in the time frame of the panel's activities (NICHD 2013a, 2013c, and 2014b). In searching for additional information about the NCS on the study's website, the panel found detailed research results only in white papers dated 2004 and 2005, which the website states are not a current, accurate representation of the plans for the NCS and are provided for historical purposes only.

TABLE A-1 Information Provided and Not Provided by NCS Program Office in Response to Panel Requests

Topic	Information Provided and Source	Information Not Provided		
		Requested Information	Source	Explanation of Response
Sample Size and Design	General description: NICHD (2013b, pp. 16-17, 20-28, 40-41) and NICHD (2013d, pp. 11-16, 42-53)	(1) Number of hospitals, providers, women to be sampled at each phase; (2) costs and variances	NICHD (2013d, p. 51) and NICHD (2013h, p. 1)	Response said requested information not available because sample design still under development.
	Useful detail about recruitment process and success at various stages in Vanguard Study, including assessment of sample representativeness: NICHD (2013b, pp. 26-28), NICHD (2013d, pp. 5-10), NICHD (2013e, pp. 1-3), and NICHD (2013f, pp. 3-4)	Rationale for moving from area-based sample to hospital-based sample	NICHD (2013d, pp. 52, 57) and NICHD (2013g, p. 4)	First response said hospital approach will minimize bias and save resources. Second response cited greater operational flexibility, easier to find a replacement should a sampled hospital decline. No additional details provided.
		Details of planned reconciliation of data when each hospital has a different 4-year birth period	NICHD (2013d, p. 50) and NICHD (2013h, p. 3)	First response said NCS will do reconciliation. Second response had no reference to reconciliation. No additional details provided.

		(1) Power calculations for specific exposures and prevalence to justify sample size; (2) precision of key estimates needed to determine sample size	NICHD (2013d, pp. 43-46, 50); for (1) reformulated in NICHD (2013h, p. 2); for (2), reformulated in NICHD (2013g, p. 2)	(1) Response said any proposed exposure or outcome with defined prevalence fits within the framework used by NCS, but calculations not available because PSU sample design not complete. Second response stated that this sample size (50,000 births) "is adequate for hypotheses that we and others have proposed to make exposure outcome associations." (2) Response said NCS should be general purpose.
Hospitals as PSUs	General description: NICHD (2013b, pp. 22-25) and NICHD (2013d, pp. 11-13, 16, 54-58) Brief summary of NCS consultants' initial report: NICHD (2013i, pp. 1-2) Summary of Vanguard Study experience recruiting women at hospitals: NICHD (2013f, pp. 2-3) and NICHD (2013g, p. 2)	Details of proposed sample of hospitals including quality of the frame.	NICHD (2013d, pp. 54-58), NICHD (2013f, p. 1), NICHD (2013h, p.1) and NICHD (2013j, p. 3)	Response said details not available because sample design under development.

continued

TABLE A-1 Continued

Topic	Information Provided and Source	Information Not Provided		
		Requested Information	Source	Explanation of Response
Prenatal Providers as SSUs	General description: NICHD (2013b, pp. 20-24) and NICHD (2013d, pp. 13-15, 70-74)	Detail about quality of list frame of providers	NICHD (2013d, pp. 70-71) and NICHD (2013h, p. 1)	First response described construction of list. Quality not mentioned. Second response provided more detail on construction of list.
	Summary of Vanguard Study experience sampling and recruiting both providers and women at providers: NICHD (2013f, pp. 3-4) and NICHD (2013g, p. 1)			
Size and Cost of Birth and Prenatal Strata	General description: NICHD (2013b, pp. 24-25), NICHD (2013d, pp. 12-16, 59-63)	Cost analysis that was used to justify large birth stratum	NICHD (2013d, pp. 59-60)	Response said a subsequent review of submitted cost data is suggesting a need to reevaluate this cost estimate; no additional details provided.
	Some cost information provided verbally to panel's cost consultants and in NICHD (2013j)	Detailed cost information, including cost metrics for Vanguard Study	NICHD (2013d, pp. 59-63), NICHD (2013h, pp. 1-3), and NICHD (2013j, p. 2)	Many details not provided.
	Reviewed and commented on panel's cost appendix: NICHD (2014a)			

Convenience Samples	General description: NICHD (2013b, pp. 25-26) and NICHD (2013d, pp. 16-18, 64-66) Sample of nulliparous women: NICHD (2013f, p. 5) and NICHD (2013g, pp. 10-11)	Detail about composition of 5,000-child convenience samples (other than nulliparous women)	NICHD (2013d, pp. 64-66) and NICHD (2013h, p. 3)	First response said convenience sample composition not defined until sample design complete and PSUs selected. Second response said unique exposures sample only one of multiple options.
Sibling Births	General description: NICHD (2013b, p. 17) and NICHD (2013d, pp. 19-24, 67-69)	Methodology for computing selection probabilities of sibling births	NICHD (2013d, p. 68) and NICHD (2013f, p. 4)	Response said siblings included with certainty. No additional details provided.
Study Visit Schedule	General description: NICHD (2013b, pp. 32-33) and NICHD (2013d, pp. 35-41, 75-77)	Rationale behind selection of specific observation points and data to be collected	NICHD 2013d, p. 75) and NICHD (2013h, p. 2)	Response cited need for frequent measurements during periods of rapid change; no additional details provided.

continued

TABLE A-1 Continued

Topic	Information Provided and Source	Information Not Provided		
		Requested Information	Source	Explanation of Response
Approach to Measurement	Health phenotype concept: NICHD (2013b, pp. 29-32) and NICHD (2013d, pp. 26-34) General description of content and content development process: NICHD (2013b, pp. 34-35, 42-47) and NICHD (2013d, pp. 25-32, 34-35, 89-99)	(1) Information that spans early waves of interviewing, specifically about content, timing, and sample allocation details; (2) process used to select domains and subdomains and related measures (Vanguard study has data collection experience up to 12 months of age); (3) operationalization of this process; (4) measures used and how selected	NICHD (2013d, pp. 46-48 and 87-99); reformulated in NICHD (2013g, pp. 7-10)	(1) Response referred to general background section of report. Appendix states "specified visits include what is currently being tested/planned in the Vanguard Study and are possible measures to be collected in the Main Study." Table shows only categories of measures (environmental sample, biospecimen, questionnaire.) (2) Response stated Vanguard Study not designed to examine prevalence of exposures or outcomes or examine exposure-outcome relationships. (3) Response stated that within each domain and subdomain measures were selected to address specific topics. (4) Measures not provided.
	Exposure assessment: (1) summary of earlier exposure assessment workshops:[1] NICHD (2013b, pp. 48-51); (2) summary of Vanguard Study research	(1) Clarification of information collected for birth stratum, and, if retrospective assessment is an objective, provide specific documentation	NICHD (2013d, pp. 43-44, 48, 76), and for (1), NICHD (2013g, p. 10)	(1) Response states NCS is exploring the option of self-collected environmental samples for the birth stratum, and that the purpose of a postbirth visit is to collect information about persistent environmental.

	in environmental collections: NICHD (2013g, pp. 8-10)	of validity of methods for the full spectrum of environmental factors. (2) Besides environmental exposures (dust, etc.) what other measures of the family environment are considered important. (3) Strategy for nonpersistent exposures		exposures. No additional information about retrospective assessment provided. (2) Response states that NCS will capture a full spectrum of environmental and social factors. Details not provided. (3) Nonpersistent exposures would rely on probability and timing to detect.
Health Disparities	General description: NICHD (2013b, pp. 11-12, 19-20, 26-28) and NICHD (2013d, pp. 78-81) Useful detail about recruitment by demographic category from Vanguard Study: NICHD (2013d, pp. 5-10) and NICHD (2013e, pp. 1-3).	Panel requests NCS reaction to their judgment that oversampling within the context of the main study is the way to best address health disparities	NICHD (2013d, pp. 78-81), and NICHD (2013g, p. 11)	Response states that choosing groups to oversample is challenging and oversampling reduces the precision of estimates of the overall population.

continued

TABLE A-1 Continued

Topic	Information Provided and Source	Information Not Provided		
		Requested Information	Source	Explanation of Response
Issues Related to Scientific Merit and Generalizability		NCS staff qualifications	NICHD (2013e, p. 5)	No response provided.
		Schedule for launching main study	NICHD (2014e, p. 6)	Main study field work anticipated to begin in fiscal year 2015; no additional details provided.

NOTES: PSUs = primary sampling units; SSUs = secondary sampling units.

[1]Workshop sponsored by National Children's Study, U.S. Environmental Protection Agency, and National Institute of Environmental Health Sciences.

Appendix B

Field Costs for the National Children's Study: First 7 Years

The panel was charged with evaluating the proposed methodologies for the National Children's Study (NCS) Main Study, including whether such methodologies are likely to produce scientifically sound results. In trying to carry out this part of our charge, the panel found that the relevant cost information was not available (see Chapter 5). Thus, in order to assess alternative approaches with key design elements, such as sampling frame and design, the recruitment and retention process, and broad aspects of the interview schedule and data collection, the panel commissioned two consultants—Lisa Schwartz of Mathematica Policy Research and Randall Olsen of Ohio State University—to construct a cost model of several design options. They collected cost information relevant to the modeling at meetings and through telephone conversations with staff from the NCS Program Office. As explained below, most of the NCS Program Office cost suggestions were adopted and used in this analysis, and the Program Office staff were invited to "fact-check" the model assumptions and results. Key results of the cost analysis are summarized in Chapter 5. This appendix details the assumptions behind the cost model.

SCOPE OF THE COST ANALYSIS

The assessment was limited to the field costs that would vary to a material degree for the design options the panel considered. Specifically, the budget models focused on recruitment, retention, and data collection costs through the first 7 years of the Main Study. Our general objective was to assess the cost

implications of key design tradeoffs that were roughly consistent with the total estimated field costs of the current proposed Main Study design.

Despite the large 7-year total field costs we estimate—some $1.3 to $1.6 billion—it is important to realize that our estimates omit many potentially sizable study costs, although they are not materially affected by total sample size or recruitment strategy. Those omitted costs include

- project management and oversight of contractors by the National Institute of Child Health and Development (NICHD);
- any costs associated with the Vanguard Study and pretesting at any sites;
- any costs associated with purchasing or building the sample frame;
- design and supervision of the sampling activities and the production of sample weights;
- design and review of questionnaire and specimen collection protocols and the preparation, printing and distribution of training materials or other supplies and consumables to the field;
- costs to assay environmental samples or analyze blood samples other than what the Program Office determines is necessary for quality control;
- security tasks to keep the NCS compliant with the Federal Information Security Management Act of 2002;
- management of the data collection contractors;
- management of the data flow to and from the field to whichever contractor is handling the data;
- programming of the questionnaires, configuration and distribution of laptops, including troubleshooting repair, replacement, and inventory control;
- tracking and storing all specimens; and
- data management, including storage, data cleaning, variable creation, geocoding, appending of other data to the master database, documentation, generation of user data files whether for public access or restricted access, user support, user outreach of any kind, and publicity and promotion of the data.

MODELING ASSUMPTIONS

To the extent possible, the panel constructed a cost model based on information provided by the NCS Program Office. When cost information was not available, we relied on the experience of our consultants (who have a combined 40 years of experience with survey research and several contract survey research firms) to develop reasonable estimates. The cost models reflect the breadth of

their experience in budgeting various components of large, complex data collection efforts.

For the most part, the Program Office agreed with our cost assumptions. When their review suggested changes to our assumptions (either up or down), we usually revised our models accordingly. These changes to our initial assumptions did not produce major changes to the overall estimates. As described in the penultimate section of this appendix, we questioned only two NCS Program Office assumptions as being inconsistent with the experience of the consultants who conducted the cost analysis. In one case, the need for some in-person interviews at birth, the consultant's assumption was used in the cost analysis.

We took as our baseline model the design as described by the NCS Program Office in documents submitted to the panel.[1] We begin with a high-level summary of the designs included in our cost models. Across all models, we assume the yield from provider-based sample recruiting is measured by the number of prenatal enrollments that lead to a sampled birth. That is, if a woman agrees to the study and makes an appointment for the first prenatal interview, her agreement is not counted as a successful recruitment until and unless the woman completes the first prenatal interview and then stays with the study through childbirth.

Based on the information provided to the panel, we assume that 250 hospitals are the primary sampling units (PSUs).[2] Also based on Program Office information, we assume that 50 percent of women who are approached will refuse to participate and that of the women who agree to participate, 20 percent will drop out of the study after their first prenatal interview and before birth. The Program Office estimates the average enrolled prenatal birth mother will receive 1.5 prenatal interviews. Accounting for attrition, this means about 1.7 prenatal visits are conducted for every birth brought into the study by prenatal provider-based recruitment.

The secondary sampling units are non-hospital-based providers (i.e., practices), and we assume, on average, five such providers are recruited per sampled hospital (N = 1,250 providers). Recruitment of hospitals and providers and sample enrollment occurs in four sequential groups of PSUs. Each group of PSUs is tracked for 4 years, but the initial year for each group is staggered: that is, recruitment in group A takes place in years 1 through 4,

[1]Our work is based on materials received from the Program Office through January 2014, but we note that there may have been subsequent changes to this design, which would not be reflected in our cost assessments. However, the major cost drivers are parameterized in the spreadsheet, which is available at http://sites.nationalacademies.org/DBASSE/CNSTAT/National_Childrens_Study_2014/index.htm [July 2014]. .

[2]At the time of this report, NCS anticipated that between 200 and 300 hospitals would comprise the study's primary sampling units and estimated that they would recruit five providers per hospital.

group B in years 2 through 5 and so forth.[3] This approach spreads out the higher costs of enrollment and the heavy first-year data collection over more years so that peak costs are not as high as they otherwise would be. Using this approach, peak costs occur in year 4 when all PSUs are involved in active recruitment or data collection. The last birth would be enrolled in year 7 of the data collection phase: hence, our analysis is budgeted over 7 years. In years 8 and later, the data collection costs we estimate fall until they reach a steady state of about $50 million (plus an adjustment for inflation). This approach reflects a case load averaging 50,000 active cases per year[4] minus attrition.

We note that for modeling purposes, recruitment costs are influenced by which of two data collection protocols is used for a case. As noted above, the prenatal protocol, the more costly of the two, includes, on average, 1.5 prenatal in-home interviews, which include environmental and biospecimen collection. All women recruited prenatally and all women in the preconception sample once pregnancy is confirmed receive this protocol. The protocol for women recruited at delivery includes a take-home kit for collecting and shipping environmental samples: in all other ways the women recruited after delivery receive the same schedule of interviews. All women recruited at the hospital and women in the non-probability (non-preconception) sample receive the protocol for women recruited at delivery.

The scientific value of the prenatal sample led the panel to consider options for increasing the prenatal fraction of the sample without increasing the total field costs, which could be done in a variety of ways. As we detail below, dropping the preconception sample and then making 97 percent of the sample prenatal through provider-based recruitment comes very close to achieving the goal. To match the 7-year cost of the baseline model, the sample only has to be reduced to about 96,000. This result reflects the high cost of the preconception sample, primarily due to the large number of women who must be tracked to achieve 5,000 enrolled births.

The Cost Models

The panel's cost analysis is based on eight models, of which the first is the baseline:

[3]At the time of this report, NCS anticipated that provider-based recruitment will occur over a 3-year period in order to ensure that children born to enrolled women are born in year 4 of recruitment; delivery-based recruitment will occur over 4 years.

[4]After age 5, children will be monitored once every 2 years. Thus, to monitor 100,000 children, there will be 50,000 cases per year.

- **Model 1–baseline (current design):** 100,000 births (including siblings), with 90,000 probability sample (half recruited prenatally, half at birth), 5,000 preconception and 5,000 other convenience supplemental sample. Of the 100,000 births, 5,000 are recruited preconception and, if the mothers become pregnant within the 4-year window for recruitment, they also receive the prenatal sample protocol. The baseline protocol administers an average of 1.5 prenatal interviews for the women sampled prenatally. NCS estimates a 20 percent attrition rate between the initial prenatal, face-to-face interview and delivery, so the effect on costs is equivalent to an additional 0.2 prenatal interviews (for a total of 1.7) administered per woman enrolled in the prenatal protocol. The baseline model budget includes performing the full complement of assessments in the first year: a hospital interview and, for 15 percent of the sample, an in-home visit immediately after birth; a telephone interview at 3 months and 9 months; and in-home assessments at 6 months and 1 year. The last two assessments are performed for all birth events. The content of the various interviews may differ for the prenatal and birth sample mothers, but we assume these differences are not material in terms of overall costs.
- **Model 2–maximum prenatal recruitment:** Same as the baseline except 97 percent of the births, including those in the other supplemental convenience sample, are enrolled prenatally. We assume that in order to sample a pro rata share of women who are not eligible for prenatal, provider-based recruitment, 3 percent must be enrolled at delivery.
- **Model 3–maximum prenatal recruitment but no preconception first-birth supplement sample:** Same as Model 2 except no preconception cohort is recruited, and the probability sample is increased by 5,000.
- **Model 4–drop one in-home and telephone interview:** Same as Model 3 except eliminate one face-to-face and one telephone interview in the child's first year of life.
- **Model 5–drop one telephone interview:** Same as Model 3 except eliminate one telephone interview in the child's first year of life.
- **Model 6–drop one face-to-face interview:** Same as Model 3 except eliminate one face-to-face interview in the child's first year of life.
- **Model 7–just one prenatal interview:** Same as Model 3 except women have only one prenatal interview prior to birth.
- **Model 8–maximum prenatal recruitment, no preconception supplemental sample, no other convenience sample, and cost neutrality:** Same as Model 3 except there are no convenience samples and only 96,000 children are enrolled in the study.

Cost Model Assumptions

In this section, we provide information on some of the key cost parameters for the NCS field work, in five areas: general, sample recruitment and retention, interviewer recruitment and training, data collection, and specimen processing.

General Assumptions

- Inflation is estimated at 2 percent per year with field work beginning in 2015. Insofar as field work begins after 2015, all numbers should be increased by 2 percent per year. Should inflation accelerate, figures would need adjustment.
- Attrition is based on how long the mother has been in the study, not the age of the child. In families that include siblings, attrition by the mother results in attrition for all her children. We assume in the second year, 90 percent of mothers remain in the study; in year 3, 87 percent; in year 4, 85 percent and then a 1 percent decline in participation rate every year thereafter through year 7.
- All births to sampled women inside the 4-year recruitment period are included in the study. For a 100,000-child sample, about 8.8 percent of the children are subsequent siblings of the sampled births, based on the panel's analysis of fertility data. For sampled pregnancies occurring later in the cluster's 4-year recruitment window, there is less time for eligible siblings to be born within the 4-year window. As a result, the 8.8 percent subsequent sibling projection may be lower than one might intuitively expect.
- Indirect costs and fees are 57 percent. The fringe rate for interviewers is 20 percent and 40 percent for professional staff. Insofar as telephone interviewing is used, we assume it is done from a call center with a total, all-inclusive staff cost of $30 per hour. This assumption may be low if calling is done from a national survey firm's premises.

Sample Recruitment and Retention

- 250 hospital PSUs and 1,250 practice providers are recruited at 100 professional hours of senior level staff time per entity (150,000 professional staff hours).
- Retention of hospitals and practices in the study requires one on-site meeting per year. Hence, during the 4-year recruitment period, there would be four on-site meetings with each participating hospital and three on-site meetings with each participating practice.
- The costs of institutional review boards (IRBs) are borne by the NCS Program Office: we assume that one central IRB will approve the study

and that no additional review boards will require approval. Thus, no IRB-related costs are included in our analysis.

- Hospital staff require a stipend for collecting cord blood, cord samples, and placenta samples at birth. Estimated hourly rate for hospital staff to collect biospecimens is $100, unloaded, and it is estimated that biospecimen collection will take, on average, 0.5 hours per birth.
- 80 percent of hospitals and 70 percent of providers agree to participate, factoring in refusals to the provider recruitment model.
- To secure cooperation, a one-time monetary incentive to hospitals of $500 per hospital is included.
- A combination of monetary and nonmonetary incentives to providers is estimated at $500 per year per practice.
- To recruit the prenatal sample, field staff are on site at providers' offices for 4 hours per day, 2 days per week for a 3-year period.
- To recruit the birth sample, field staff are on site at hospitals for 20 hours per week over 4 years. On-site time does not include the hours needed to complete the in-hospital interview.
- 25 percent of the recruitment costs can be assigned to each of the four clusters of hospitals. The costs for each cluster are not distributed evenly over the 4 years of recruitment within each cluster. Rather, we assign 40 percent of the cluster costs to the first year of recruiting within a cluster, 25 percent to year 2, 20 percent to year 3 and 15 percent to year 4. This assumption reflects the up-front costs for recruiting hospitals and providers.
- The strategy for selecting the convenience samples has not been defined with enough exactitude to generate reliable budget estimates. We have assumed NCS can obtain this supplementary sample at roughly the same unit cost as either the prenatal or delivery sample.
- The participation rate among recruited women is 50 percent, requiring a sample of about 11,650 per year for each of the four major cluster groups to meet the overall target sample size.

Interviewer Recruitment and Training

- Interviewer recruitment is 30 hours of a field supervisor's time per recruit, which includes time to recruit interviewers who will do the phlebotomies.
- Training is 80 hours, including time to train for phlebotomy, and requires 15 days of hotel and per diem at $300 per day. Travel to training sites is projected at $600 per trainee.
- Telephone interviewer training is projected at 40 hours and does not require travel by trainees. We assume this training will be done in person at the call center.

Data Collection

- All births include an in-hospital interview of the mother and the collection of the placenta and cord blood and maternal urine and saliva, blood, and dried blood spots from the infant.
- 15 percent of births will need an age 0 in-home interview to collect environmental samples, biospecimens (blood and breast milk, at a minimum) that respondents do not complete and return themselves.
- One field interviewer is sent to households to complete an in-person interview and specimen collection.
- 65 percent of all in-person cases are "easy" and require 5.5 hours per respondent interviewed for contact activities (including travel, appointment setting, rescheduling, and packing specimens).
- 25 percent of all in-person cases are "hard" and require 9 hours per respondent interviewed for contact activities.[5]
- In-person interviewing in the home is 2 hours per interview including consent, answering respondent questions and maintaining rapport, anthropometrics, and specimen collection. This time is in addition to the time required to locate, make contact, travel, and secure cooperation with a respondent.
- All in-person interviews include a venous blood draw from the mother or child.
- Interviewer travel expenses $40 to complete an easy case and $110 for a hard case.
- Respondent fees average $75.
- Post-birth telephone interviews are 1 hour per interview. Telephone interviewers require 2 hours to secure cooperation for easy cases and 4.5 hours for hard cases. Hard telephone cases will also require, on average, $50 each in other expenses and some will require a personal visit to secure cooperation.[6]
- Laptops and smartphones, including connection and Internet charges are $750 per year per interviewer.
- Anthropometric equipment (tapes, scales, calibration weights, stadiometers blood pressure meters, etc.) is $4,000 per interviewer in year 1, $250 annually thereafter, based on experience in the Vanguard Study. We assume a 3-year lifetime for this equipment, making the

[5]Difficult cases often require repeated visits due to broken appointments, as well as visits to determine where the respondent is located and to secure additional locating information from former neighbors. Younger households are often more mobile, which makes some cases difficult because of locating problems and not just because of reluctance or resistance to cooperation. The cases that are not "easy" or "hard" are noninterviews.

[6]Other expenses include additional mailings and an in-house locating effort, such as electronic database searches, to obtain new contact information.

annual cost $1,550 per interviewer. The cost of Oragene kits and specimen supplies is $50 per case. The cost for collecting, packing, mailing, labeling and assaying 5 percent of the environmental and blood samples[7] is $190 per in-person interview.

- The expected interview completion rate is 90 percent.
- Field management is estimated at 1 hour of field supervisor time for every 5 hours of field interviewer time, and 1 hour of higher level supervision for every 6 hours of field supervisor time.
- Field interviewers' rate of pay is $23 per hour to compensate for the specialized skills that are needed to do phlebotomy; field supervisor pay is $35 per hour; and the next higher level of supervision is set at $40 per hour. We do not budget overall field management and central direction as the scope for such work is unspecified as yet.

Specimen Processing

- No lab work will be done on the environmental samples; lab work will be done on 5 percent of the venous blood samples for quality control only.
- Shipping costs for specimens from the field to the central repository are approximately $135,000 per year per 1,000 participants.

Our assumptions do not include the costs of repository processing and storage of biological and environmental samples, which we estimate at $126 million over the first 7 years ($18 million in annualized costs), based on an inventory of 42,000,000 primary and derivative samples. We note these for informational purposes only as they are significant.

COSTS AND COST DRIVERS

Based on these assumptions and the resources implied for the NCS design, we estimated the costs of our models. We begin this section with the major factors that influence the costs.

- Sample size—Pursuing a cohort of 100,000 is the primary driver of costs. Even modest cost perturbations at the interview level, given the number of planned interviews and the sample size, will generate large impacts in dollar amounts, if not in percentage terms. A corollary to this fact is that if large cost adjustments are necessary, sample size is a necessary dimension to include.

[7]NCS plans to assay or analyze 5 percent of environmental and biological samples for quality control purposes only.

- Recruitment at hospitals and providers—The cost of recruiting the probability sample is roughly $231 million or roughly 15 percent of the 7-year costs of the project. This cost varies with the fraction of the sample enrolled at delivery rather than prenatally.
- Interviewing costs—At roughly $1,000 each (in current dollars), adding or dropping one face-to-face interview for each respondent changes costs by about $100 million. Changing the number of telephone interviews per respondent by one changes total costs by about $40 million. As noted above, we assume that to obtain a satisfactory rate of completed remote interviews, some respondents will need to be interviewed or at least contacted in person to secure an interview.
- Preconception sampling—Because several women (about 3.6 based on NCS program office estimates) need to be tracked in order to yield one woman with a birth entering the sample, the cost of obtaining 5,000 preconception births enrolled in the study is about $76 million.
- Fraction in the prenatal protocol group—Changing respondents from being recruited at providers and administered prenatal interviews rather than being recruited at birth and not receiving prenatal interviews increases project costs by $30 million for every 10 percentage point increase in the fraction receiving the prenatal protocol. This cost reflects both the cost of recruitment and the cost of prenatal interviews, both to women who enroll in the study and the women who attrite before delivery.
- Interview length—Reducing by 5 minutes the length of every interview administered in the first 7 years of the study, both face to face and telephone, reduces costs over the first 7 years by $10 million. Attempting to save money by reducing interview length is relatively ineffective and will come at far higher price, scientifically, than eliminating an entire interview. We believe this is the least attractive approach to cost containment of all the alternatives we explored.

For the baseline and models 2 through 8, Table B-1 shows our estimated field costs for the first 7 years of the NCS, the cost differential relative to the baseline, and the sample size that would make each model cost neutral with respect to the baseline. (This table also appears as Table 5-1 in Chapter 5). A notable feature of the table is that while moving to an overwhelmingly prenatal protocol for data collection increases costs, those additional costs are substantially offset by eliminating the preconception cohort. Modest reductions in the schedule of interviews or modest reductions in total sample size (or both) result in costs that are roughly equivalent to baseline costs. These alterations make it possible to collect a probability sample of births that is larger than the baseline proposal. Other adjustments are possible, but this provides a framework for looking at costs and reconciling them with budgets.

TABLE B-1 Field Costs of Alternative Sample Designs

Model and Sample Size*	Total Field Cost (in millions of $)	Base Field Cost–Modeled Field Cost (in millions of $)	Size of Main Sample (including siblings) for Field Cost Neutrality
N = 100,000			
Baseline: current design—one-half prenatal and one-half birth recruitment for 90,000 births, plus 5,000 preconception supplemental sample and 5,000 convenience sample (Model 1)	$1,495	NA	NA
Maximum prenatal recruitment plus preconception supplemental sample, and convenience sample (Model 2)	$1,631	($135)	89,179
Maximum prenatal recruitment and convenience sample but no preconception supplemental sample (Model 3)	$1,542	($47)	96,256
Drop one in-home and one telephone interview from Model 3 (Model 4)	$1,348	$147	113,940
Drop one telephone interview from Model 3 (Model 5)	$1,492	$3	100,256
Drop one in-home interview from Model 3 (Model 6)	$1,398	$98	108,836
Have only one prenatal interview in Model 3 (Model 7)	$1,488	$7	100,600
N = 96,000			
Maximum prenatal recruitment, no preconception or other supplemental sample and cost neutrality (Model 8)	$1,495	($0)	96,256

NOTES: All cost figures are based only on costs of fielding the study and thus do not represent the full cost of the NCS; see text for discussion. Cost of convenience sample assumed to be equal to cost of equal-sized prenatal sample.

*See text for details of the panel's models.

TABLE B-2 Year-by-Year Cost Comparisons Across Models (in millions of $)

Year	Baseline	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
1	$ 62	$ 70	$ 62	$ 53	$ 60	$ 56	$ 59	$ 61
2	127	142	129	110	124	115	122	126
3	206	230	212	180	204	188	201	206
4	307	340	317	271	306	283	304	308
5	294	320	306	266	296	276	295	296
6	272	291	281	251	274	259	275	272
7	227	238	234	216	229	221	233	226
Total	1,495	1,631	1,542	1,348	1,492	1,398	1,488	1,495

NOTE: See text for discussion.

Table B-2 shows the estimated yearly costs for the eight models as well as the 7-year total. The data show that, despite spreading sample recruitment costs over 7 years instead of four, field costs peak at a level above the fiscal 2014 appropriation level. Attempting to spread recruitment costs over even more years may reduce the peak but increase total program costs.

Questionable Assumptions

The panel questioned only two Program Office assumptions as being inconsistent with the experience of the consultants who conducted the cost analysis for us, recruiting costs and need for follow-up. First, the Program Office estimates that the level of effort needed to recruit for the Main Study will be the same as that required for the Vanguard Study. However, the original Vanguard Study contractors were primarily universities, many of which also had medical schools and university-affiliated hospitals. Staff affiliated with those universities or contractors acting on their behalf recruited hospitals, providers, and women for the study. In our judgment, these affiliations and the universities' community-level "brand equity" likely had two effects: (1) they made it more likely that the entities they approached agreed to participate; and (2) the level of effort needed to gain that cooperation was less than it would otherwise be. We judged that it is unlikely that a contract research organization will be as effective in garnering support from a random sample of hospitals, providers, and women as an effort undertaken by a university or medical center with strong national or regional name recognition. While all of our models incorporate the Program Office's recruiting assumptions, we note in Chapter 5 that the study may incur substantially higher costs if, in practice, contract research firms must expend more effort than estimated to achieve the study's recruitment goals. A number of changes to recruiting assumptions could result from this change in staffing model:

- In addition to 100 hours of professional staff time needed to recruit hospitals, up to 300 additional hours may be needed to gain endorsement from professional organizations (such as the American Hospital Association); a similar number of hours of professional staff time may be needed to gain support from professional organizations that can encourage providers to participate in the study.
- Two rather than one on-site meeting may be needed per year per hospital and per provider to recruit and retain them in the study; in addition, monthly phone meetings may be needed to address issues and maintain relationships.
- Contract research firms may host webinars for participating practices to encourage the sharing of best practices (no webinars were included in the baseline model).

- Practices may require a substantially higher incentive than the budgeted $500 per year for hospitals to participate: our baseline model assumes a one-time payment of $5,000 for practices to secure participation.
- Field staff may need to be on site for more hours each week to successfully recruit women for the study. For the delivery sample, we assumed field interviewers would need to be on site 33 hours per week (the baseline model assumed 20 hours per week); for the provider-based sample, we assumed field staff would be on site 16 hours per week (the baseline model assumed 8 hours per week).

Second, in all of our models including the baseline we assumed 15 percent of cases would need a follow-up visit after birth to retrieve specimens or conduct an interview. The Program Office assumes this would not be necessary, but we judge that the many distractions surrounding a birth, the possibility of rapid discharge from the hospital before the in-hospital interview is conducted, and the possibility some mothers simply would not get around to collecting environmental samples all suggested we allow for the need for an in-home visit shortly after birth.

Limitations of Our Cost Model

Apart from limiting its estimates to only the field costs of the NCS, our cost analysis has other noteworthy limitations. We believe our estimates are reasonable, but there are reasons that contractors may bid costs that differ from those presented here. For example, if the contracts are fixed-price rather than cost–reimbursement contracts, the uncertainties inherent in performing such a large and complex project may lead bidders to build in a "cushion" to cover uncertainties, such as subtle changes in the interpretation of the scope of work, delays created by procurement problems, or other administrative or governmentwide disruptions. If the work scope is not clearly defined due to uncertainties over specimen collection or interview content, contract modifications become more likely, and such changes are inherently less competitive, increasing costs. There is also the risk of a contractor winning a bid that is based on a low tender price and then being unable to perform the work within the dollar limits or finds the work technically more demanding than expected. Such a problem with one part of the project, say, technological infrastructure, can easily affect other parts. Subtle changes to questionnaires can, depending on the change, have serious cost implications that only become evident when they are fielded.

There are also risks inherent in survey projects for which respondent reaction and attitudes have unpredictable cost implications. For example, the plan to draw blood from the mother at every personal visit is a very aggressive plan in terms of the demands it places on respondents. This may lead to a less

cooperative respondent pool that may require more time, effort, and incentive payments to overcome. Alternatively, these demands may reduce costs because respondents leave the study in greater numbers. There are also regulatory risks. If regulations to which the NCS is subject become more stringent, this could easily generate unpredictable cost increases. These kinds of problems are not specific to the NCS, of course.

Appendix C

Biographical Sketches of Panel Members and Staff

GREG J. DUNCAN (*Chair*) is distinguished professor of education at the University of California, Irvine. Previously, he was a professor at the University of Michigan and director of the Panel Study of Income Dynamics. He was a member of the 2008 Panel to Review the National Children's Study Research Plan. His recent work has focused on estimating the role of school-entry skills and behaviors on later school achievement and attainment and the effects of increasing income inequality on schools and children's life chances. He has served as president of the Population Association of America and of the Society for Research in Child Development. He received the 2013 Klaus J. Jacobs Research Prize of the Jacobs Foundation, given for scientific work of high social relevance to the personality development of children and young people. He is a member of the National Academy of Sciences. He has a Ph.D. in economics from the University of Michigan.

DEAN B. BAKER is professor of medicine, pediatrics, and epidemiology in the School of Medicine and director of the Center for Occupational and Environmental Health at the University of California at Irvine. He was a former principal investigator at one of the Vanguard Study locations of the National Children's Study. His research interests include environmental epidemiology, with, an emphasis on children's environmental health, including developmental toxicity associated with exposures to heavy metals and pesticides, and environmental factors in asthma; and occupational epidemiology with an emphasis on role of work organization and stress in the etiology of cardiovascular disease. He is a former president of the International Society for Environmental Epi-

demiology. He is a recipient of the Kehoe Award for Excellence in Education and Research by the American College of Occupational and Environmental Medicine. He has an M.D. from the University of California at San Diego and a M.P.H. in epidemiology from the University of California at Berkeley.

PAUL P. BIEMER is a distinguished fellow in statistics at RTI International and associate director for survey research and director of the certificate program in survey methodology at the Odum Institute for Research in Social Sciences at the University of North Carolina at Chapel Hill. He also directs the Center of Excellence in Complex Data Analysis at RTI. His work focuses on survey design and analysis, general survey methodology, and nonsampling error modeling and evaluation. Among his honors, he has received the H.O. Hartley Award, the Morris Hansen Award, and the Roger Herriot Award for excellence in statistical research. He is an elected fellow of the American Statistical Association, the American Association for the Advancement of Science, and the International Statistics Institute. He has a Ph.D. in statistics from Texas A&M University.

BARBARA LEPIDUS CARLSON is associate director of statistics at Mathematica Policy Research. Her work focuses on both sampling and survey direction, predominantly in the fields of health care and early childhood education, including a number of studies of Head Start. Her work focuses on sample design and implementation, creating sampling and analysis weights, calculating response rates, estimating design effects, and ensuring overall data quality. In addition, she is responsible for producing technical documentation of the methodology using language that is understandable by nonstatisticians. She has also directed a number of survey projects at Mathematica, including several rounds of the Community Tracking Study Household Survey (now known as the Health Tracking Household Survey) and the Evaluation of the Cash and Counseling Demonstration. She has an M.A. in mathematics/statistics from Boston University.

ANA V. DIEZ-ROUX is dean of the Drexel University School of Public Health. Previously, she held several positions at the School of Public Health of the University of Michigan, including professor and chair of epidemiology, director of the Center for Social Epidemiology and Population Health, director of the Center for Integrative Approaches to Health Disparities, research professor in the Survey Research Center of the Institute for Social Research, and director of the Robert Wood Johnson Health and Society Scholars Program. Her research interests include social epidemiology, neighborhood health effects, cardiovascular disease epidemiology, air pollution and cardiovascular risk, multilevel analysis, racial and ethnic disparities, and systems approaches

in population health. She is a member of the Institute of Medicine. She has an M.D. from the University of Buenos Aires, Argentina, and an M.P.H. and Ph.D. from the Johns Hopkins University School of Hygiene and Public Health.

NANCY J. KIRKENDALL (*Study Director*) is a senior program officer for the Committee on National Statistics. Previously, she served as director of the Statistics and Methods Group of the Energy Information Administration (EIA) and a member of EIA's senior staff. She also served as senior mathematical statistician in the Statistical Policy Branch of the Office of Information and Regulatory Affairs of the U.S. Office of Management and Budget, serving as the desk officer for the U.S. Census Bureau and chair of the Federal Committee on Statistical Methodology. She is a fellow and past vice president of the American Statistical Association and a past president of the Washington Statistical Society. She is a recipient of the American Statistical Association's Roger Herriot Award for innovation in federal statistics and its Founder's Award. She has a Ph.D. in mathematical statistics from George Washington University.

VIRGINIA M. LESSER is professor and currently chair in the Department of Statistics at Oregon State University. She has served as the Director of the Oregon State University Survey Research Center since 1993. Her research interests include sampling, survey methodology, environmental statistics, and applied statistics. Her current research projects include investigations to compare response rates, costs, and errors in single-mode vs. multi-mode surveys. She is an elected fellow of the American Statistical Association. She has a doctorate in public health in biostatistics from the University of North Carolina.

MARIE C. McCORMICK is Sumner and Esther Feldberg professor of maternal and child health in the Department of Social and Behavioral Sciences in the Harvard School of Public Health, where she was formerly chair of the Department of Maternal and Child Health. She is also professor of pediatrics at Harvard Medical School, senior associate for academic affairs in the Department of Neonatology at the Beth Israel Deaconess Medical Center, and senior associate director of the Infant Follow-up Program at Children's Hospital. Her research involves epidemiological and health services research investigations in areas related to infant mortality and the outcomes of high-risk neonates. Her current research projects include outcomes of infants experiencing neonatal complications (such as low birth weight) and interventions to ameliorate adverse outcomes; evaluation of programs designed to improve the health of families and children; and maternal health and prematurity. She is a member of the Institute of Medicine. She has an M.D. from the Johns Hopkins Medical School and a Sc.D. from the Bloomberg School of Public Health of Johns Hopkins University.

SARA S. MCLANAHAN is the William S. Tod professor of sociology and public affairs at Princeton University. She is the founding director of the Bendheim-Thoman Center for Research on Child Wellbeing and a principal investigator of the Fragile Families and Child Wellbeing Study. She was chair of a steering committee that organized a workshop held in 2013 to review the National Children's Study research plan. Her research interests include family demography, poverty and inequality, and social policy. She is a past president of the Population Association of America and has served on the boards of the American Sociological Association, the Population Association of America, and the Institute of Medicine's Board on Children, Youth, and Families. She currently serves on the board of the Russell Sage Foundation. She was named the James S. Coleman fellow of the American Academy of Political Science, and she is a member of the National Academy of Sciences. She has a Ph.D. in sociology from the University of Texas at Austin.

GEORGE R. SAADE is professor of obstetrics and gynecology, chief of obstetrics and maternal-fetal medicine, and director of the Perinatal Research Division at the University of Texas Medical Branch at Galveston. He is also a professor of cell biology and directs a laboratory with a focus on maternal-fetal physiology and developmental programming of adult diseases. His clinical interests are in hypertensive disorders in pregnancy, preterm labor, and medical complication of pregnancy. He was a member of the National Children's Study pregnancy and infant group when it was under development. He is a past president of the Society for Maternal-Fetal Medicine. He is a fellow of the American Congress of Obstetricians and Gynecologists. He has an M.D. from the American University of Beirut.

S. LYNNE STOKES is a professor in the Department of Statistical Science at Southern Methodist University. She was previously a staff member at the U.S. Bureau of the Census, where she worked on measurement of interviewer errors for the Current Population Survey. Her current research interests include sampling and nonsampling error modeling, psychometrics, and capture-recapture methodology. She is an elected fellow of the American Statistical Association (ASA) and a recipient of ASA's Founder's Award. She is a member of the technical advisory committee for the Department of Education's National Assessment of Educational Progress. She has a Ph.D. in mathematical statistics from the University of North Carolina at Chapel Hill.

LEONARDO TRASANDE is associate professor in pediatrics and environmental medicine at the School of Medicine and in health policy at the Wagner School of Public Service, both at New York University. His research focuses on identifying the role of environmental and other factors in chronic childhood disease and documenting the economic costs for policy makers of failing to

prevent them proactively. He has focused particularly on the costs of children's exposure to chemicals in the environment, on increases in hospitalizations associated with childhood obesity, and on increases in medical expenditures associated with being obese or overweight in childhood. He serves on the Environment Programme Steering Committee of the United Nations, which is developing a global outlook on chemicals policy, and on the executive committee of the Council for Environmental Health of the American Academy of Pediatrics. He recently served on the Board of Scientific Counselors for the National Center for Environmental Health at the U.S. Centers for Disease Control and Prevention. He served on the Steering Committee of the National Children's Study from 2005-2011. He has an M.P.P. from the Harvard Kennedy School of Government and an M.D. from Harvard Medical School.

COMMITTEE ON NATIONAL STATISTICS

The Committee on National Statistics was established in 1972 at the National Academies to improve the statistical methods and information on which public policy decisions are based. The committee carries out studies, workshops, and other activities to foster better measures and fuller understanding of the economy, the environment, public health, crime, education, immigration, poverty, welfare, and other public policy issues. It also evaluates ongoing statistical programs and tracks the statistical policy and coordinating activities of the federal government, serving a unique role at the intersection of statistics and public policy. The committee's work is supported by a consortium of federal agencies through a National Science Foundation grant.

BOARD ON CHILDREN, YOUTH, AND FAMILIES

The Board on Children, Youth, and Families (BCYF) is a nongovernmental, scientific body within the National Academy of Sciences that convenes top experts from multiple disciplines to analyze the best available evidence on critical issues facing children, youth, and families today. Our ability to evaluate research simultaneously from the perspectives of the biological, behavioral, health, and social sciences allows us to shed light on innovative and influential solutions to inform the nation. Our range of methods—from rapidly convened workshops to consensus reports and forum activities—allows us to respond with the timeliness and depth required to make the largest possible impact on the health and well-being of children, youth, and their families throughout the entire lifecycle. BCYF reports provide independent analyses of the science and go through a rigorous external peer review process.