

Vital and Health Statistics

Series 1

Program and Collection Procedures

Design and Operation of the National Survey of Children's Health, 2007

*Stephen J. Blumberg, Ph.D., Centers for Disease Control and Prevention;
Erin B. Foster, M.S.T., Alicia M. Frasier, M.P.H., Jennifer Satorius, M.S.W.,
Ben J. Skalland, M.S., Kari L. Nysse-Carris, Ph.D., Heather M. Morrison, M.A., and
Sadeq R. Chowdhury, Ph.D., NORC at the University of Chicago;
Kathleen S. O'Connor, M.P.H., Centers for Disease Control and Prevention*

**DEPARTMENT OF HEALTH AND HUMAN SERVICES
Centers for Disease Control and Prevention
National Center for Health Statistics
Hyattsville, Maryland**

Acknowledgments

The Maternal and Child Health Bureau of the Health Resources and Services Administration provided funding and direction for the National Survey of Children's Health.

The survey was conducted by the Centers for Disease Control and Prevention's (CDC) National Center for Health Statistics. The project director is Marcie Cynamon. Design, production, and analytic assistance for this project were provided by Rosa Avila, Matthew Bramlett, and Julian Luke. CDC's National Center for Immunization and Respiratory Diseases graciously permitted the use of the National Immunization Survey sampling frame for this survey.

NORC at the University of Chicago and its subcontractor conducted all interviews for this project. Technical assistance was contributed by Nada Ganesh, Shannon Ten Broeck, Enyinnaya Adighibe, Erin Raasch, Tracie Carroll, Santanu Pramanik, and Steven Pedlow. Gretchen Caspary, Kirk Wolter, Kennon Copeland, and Tiffani Balok provided management support.

This report was edited by [NAME] and typeset by [NAME] of CDC's Division of Creative Services. The report's outline and some text were borrowed with permission from an earlier Vital and Health Statistics series report ("Design and Operation of the National Survey of Children's Health, 2003").

Finally, we extend our appreciation to the tens of thousands of parents and other family members who were willing to share their stories. Their efforts made this project a reality.

Suggested Citation

Blumberg SJ, Foster EB, Frasier AM, et al. Design and Operation of the National Survey of Children's Health, 2007. National Center for Health Statistics. Vital Health Stat 1. Forthcoming.

This draft has not yet been edited by CDC's Division of Creative Services. Please do not use quotations from this unedited document. It has been released in advance of its formal publication to provide background and guidelines for users of the recently released data files.

Contents

Abstract	1
Introduction	2
State and Local Area Integrated Telephone Survey Program	2
National Survey of Children's Health	3
National Survey of Adoptive Parents	3
Influenza Vaccination Module	4
Sample Design	4
The National Immunization Survey Sampling Plan	5
National Survey of Children's Health Sample Design and Allocation	5
Drawing the Sample	6
Conducting the National Survey of Children's Health Interviews	6
Questionnaire	7
Content	7
Programming the Computer-Assisted Telephone Interview	9
Pretesting	10
Interviewer Training	10
Data Collection	11
Advance Letters	11
Toll-Free Telephone Numbers	12
Selection of Sampled Child	12
Selection of Respondent	13
Informed Consent	13
Assurance of Confidentiality	13
Identification of Age-Eligible Children for the National Immunization Survey	14
Spanish-Language Interviewing	14
Interviewing in Languages Other than Spanish or English	15
Interview Length	16
Interview Breakoffs	16
Cases Pending at Close of Data Collection	16
Incentive Effort	17
Response Rates	17
Alternative Response Rates	18
Efforts to Maximize Response Rates	19
Quality Control of Interviewing	19
Data Files	20
Editing	20
Missing Data	21
Coding of Verbatim Answers into Question Responses	22
Edits to Protect Confidentiality	22
Derived Variables	27
Dummy Variables	29
Additional Data Notes	29
Procedures for Developing Sampling Weights	31
Base Weight	31
Computing Full Sample Weights from Quarterly Weights	31
First Form of Nonresponse: Unknown Household Status	32
Second Form of Nonresponse: Unknown Household Eligibility	32

Adjustment for Households with More than One Child	33
Third Form of Nonresponse: Sampled Children for Whom an Interview Is Not Completed	33
Adjustment for Multiple Telephone Landlines	33
Adjustment for Noncoverage of Children in Non-Landline Telephone Households.....	33
Raking Adjustment	34
Quality Control	35
Estimation and Hypothesis Testing.....	35
Interpretation of Weighted Estimates	35
Variables Used for Variance Estimation.....	36
Variance Estimation Using SUDAAN or STATA.....	36
Variance Estimation for Subsets of the Data	37
Weighted Frequencies, Prevalence Estimates, and Standard Errors.....	37
Guidelines for Data Use.....	38
Further Information.....	38
References.....	40
Appendix I	43
Sample Design	43
Preparation of the Sample.....	44
Obtaining Addresses for Advance Letters	46
Appendix II	47
Banked Sample	47
Appendix III.....	49
Computing Sampling Weights	49
Appendix IV.....	50
Questionnaire	50
Appendix V	51
Summary of Key Differences Between the 2003 and 2007 Questionnaires	51
Appendix VI.....	56
Summary of Questionnaire Changes during Data Collection	56
Appendix VII	59
Procedures for Assigning Poverty Status.....	59
Appendix VIII.....	61
Program Names Used for Medicaid and SCHIP Questions.....	61
Appendix IX.....	62
Letters Sent to Sampled Households	62
Appendix X.....	64
Disposition Code Frequencies and Response Rate Calculation.....	64
Appendix XI.....	65
Incentive Effort	65
Eligibility	65
Procedure for Offering Tier 1 Incentives.....	66
Procedure for Offering Tier 2 Incentives.....	67
Interview Completion Rates.....	67
Appendix XII	69
Multiple Imputation of Household Poverty Level	69
Imputation Procedures	70
Results of Modeling.....	72
Use of Multiply Imputed Values.....	72
Appendix XIII	75
Prevalence Estimates and Weighted Frequencies	75

Abstract

Objectives

This report presents the development, plan, and operation of the 2007 National Survey of Children's Health, a module of the State and Local Area Integrated Telephone Survey, conducted by the National Center for Health Statistics, Centers for Disease Control and Prevention. This survey was designed to produce national and state-specific prevalence estimates for a variety of physical, emotional, and behavioral health indicators and measures of children's experiences with the health care system. The survey also includes questions about the family (e.g., parents' health status, stress and coping behaviors, family activities) and about respondents' perceptions of the neighborhoods where their children live. Funding and direction for this survey was provided by the Maternal and Child Health Bureau, Health Resources and Services Administration.

Methods

A random-digit-dialed sample of households with children less than 18 years of age was selected from each of the 50 states and the District of Columbia. One child was randomly selected from all children in each identified household to be the subject of the survey. The respondent was a parent or guardian who knew about the child's health and health care.

Results

A total of 91,642 interviews were completed from April 2007 to July 2008. Nearly 80% of the interviews were completed in 2007. Interviews were completed in 66.0% of identified households with children. The weighted overall response rate was 46.7%. A data file has been released that contains demographic information on the selected child, substantive health and well-being data for the child and his/her family, and sampling weights. Estimates based on the sampling weights generalize to the noninstitutionalized population of children in each state and nationwide.

Keywords

Child health services, chronic disease, family functioning, health status indicators, health surveys, medical home, needs assessment, pediatrics

Introduction

For nearly a century, the Maternal and Child Health Bureau (MCHB) of the Health Resources and Services Administration (HRSA) has been charged with the primary responsibility for promoting and improving the health of the nation's mothers and children. The mission of MCHB is to ensure the continued improvement in the health, safety, and well-being of America's women, infants, children, adolescents, and their families (1,2).

MCHB relies on data from population-based systems to evaluate progress toward its mission. National level data on child health and well-being are available from a number of ongoing surveys. However, valid and reliable state-level statistical estimates generally cannot be made from these national datasets for all states. Recognizing the need for health and well-being data that could be meaningfully compared across states and nationally for all children, MCHB sponsored the first National Survey of Children's Health (NSCH) in 2003 through the State and Local Area Integrated Telephone Survey (SLAITS) program. To continue monitoring the health of the nation's children, MCHB sponsored the second administration of the NSCH in 2007. This report documents the 2007 design and procedures.

State and Local Area Integrated Telephone Survey Program

The SLAITS program, conducted by the Centers for Disease Control and Prevention's (CDC) National Center for Health Statistics (NCHS), is a broad-based, ongoing surveillance system available at the state and local level to track and monitor the health and well-being of children and adults. These surveys use the sampling frame of the CDC's National Immunization Survey (NIS) and immediately follow the NIS in selected households, using its sample for efficiency and economy. NIS is a large-scale random-digit-dialed (RDD) telephone survey that uses computer-assisted telephone interview (CATI) technology to contact over 1 million households each year, screen for the presence of young children in selected households, and collect immunization history information for eligible children. The process to identify this large number of households – most of which are ultimately age-ineligible for the NIS – offers an opportunity to administer other surveys on a range of health- and welfare-related topics in an operationally seamless, cost-effective, and statistically sound manner.

Surveys conducted as part of the SLAITS system vary in content, duration, and sample size based on the research needs of their sponsors. Sponsors work with NCHS to establish parameters including sample size, questionnaire design, and other survey requirements. Since 2005, NORC at the University of Chicago has administered all aspects of survey operations. Their staff, in conjunction with SLAITS staff from NCHS, develops and tests the CATI instrument; recruits and trains interviewers, completes the targeted number of interviews, and prepares data files and final documentation. NCHS is responsible for all aspects of SLAITS administration.

SLAITS began in 1997 with a pilot test in two states, Iowa and Washington. This pilot survey included a series of questions on health, including issues of access to care, health status, and insurance (3). In 1998, a SLAITS module on child well-being and welfare issues was implemented using three samples: a general RDD sample of children in Texas, known Medicaid program participants in Texas, and known Medicaid or MinnesotaCare participants in Minnesota (3). In 2000, SLAITS fielded the National Survey of Early Childhood Health (NSECH), which collected data on parents' perceptions of

their young children's pediatric care, and examined relationships between health promotion in the pediatric office and at home (4).

SLAITS fielded the first National Survey of Children with Special Health Care Needs in 2001 and the first National Survey of Children's Health in 2003. The National Survey of Children with Special Health Care Needs was designed to produce national and state-specific estimates of the prevalence of children with special health care needs (CSHCN), describe the types of services that they need and use, assess aspects of the system of care for CSHCN, and evaluate the impact of their needs on their families (5). The NSCH was designed to produce national and state-specific estimates of the health and well-being of all children, their families, and communities (6). These were the first SLAITS studies to take full advantage of the NIS sampling frame to produce estimates for each of the 50 states and the District of Columbia (DC). In 2003, SLAITS also fielded the National Asthma Survey, which examined the health, socioeconomic, behavioral, and environmental predictors that relate to better control of asthma (7). The National Survey of CSHCN was repeated in 2005-2006 (8), and in 2007, the NSCH was conducted for the second time.

National Survey of Children's Health

According to its vision statement, MCHB strives "for a society where children are wanted and born with optimal health, receive quality care, and are nurtured lovingly and sensitively as they mature into healthy, productive adults." MCHB also seeks to ensure that "there is equal access for all to quality health care in a supportive, culturally competent, family and community setting" (2). This effort is achieved by providing block grants that are matched by state funds (1).

The NSCH was conducted to assess how well individual states, DC, and the nation meet MCHB's strategic plan goals and national performance measures. These goals call for MCHB to provide national leadership for maternal and child health; promote an environment that supports maternal and child health; eliminate health barriers and disparities; improve the health infrastructure and systems of care; assure quality care; work with states and communities to plan and implement policies and programs to improve the social, emotional, and physical environments; and acquire the best available evidence to develop and promote guidelines and practices to assure social, emotional, and physical environments that support the health and well-being of women and children. The results from the NSCH support these goals by providing an objective basis for Federal and state program planning and evaluation efforts (9).

The NSCH content is intentionally broad, and addresses a variety of physical, emotional, and behavioral health indicators and measures of children's health experiences with the health care system. The survey includes an extensive battery of questions about the family, such as parental health, stress and coping, and family activities. The NSCH also asks respondents for their perceptions of the child's neighborhood. No other survey provides the breadth and depth of information about children, families, and neighborhoods with sample sizes sufficient for state-level analyses in every state, collected in a manner that allows comparison among states (9).

National Survey of Adoptive Parents

In 2007, the Office of the Assistant Secretary for Planning and Evaluation (ASPE) and the Administration for Children and Families (ACF), both of the Department of Health and Human Services

(DHHS), collaborated with NCHS to develop the National Survey of Adoptive Parents (NSAP). The NSAP focused on the characteristics and needs of adopted children and their adoptive families. The survey was tailored to collect data from three types of adoptive families: those who adopted through the United States (US) foster care system, internationally, and through domestic private agencies.

Due to the relatively low prevalence rate of adoptive families in the US, surveys of this population have typically been conducted using targeted lists, which may or may not provide samples that are representative of the national population of adopted children and their adoptive families. In contrast, the large sample size of the 2007 NSCH enabled NCHS to identify a nationally representative sample of approximately 2,000 adopted children.

The NSCH identified households that contained at least one child under the age of 18 years and selected one child from that household as the subject of the detailed interview. When a selected child was identified as adopted, respondents were invited to also participate in the NSAP interview. The NSAP interview examined adoption agreements, post-adoption financial services, and post-adoption non-financial supports, as well as adoption-related measures of parent and child well-being. When combined with the NSCH data on the general health and well-being of children, their families, and their neighborhoods, the NSAP data provide researchers with previously unavailable insights into adopted children and their families. Details about the NSAP are included in a separate forthcoming NCHS report and will not be further discussed here.

Influenza Vaccination Module

In 2007, ASPE also sponsored a study to assess influenza vaccination coverage in children and characteristics of children who did and did not receive the vaccine in accordance with recommendations from the Advisory Committee on Immunization Practices. The study assessed risk factors in children for which the immunization may be indicated, as well as selected risk factors for adults in the household, as they could also indicate the need for influenza immunizations in resident children. These questions were integrated into the 2007 NSCH questionnaire. The target sample size for the influenza vaccination module was approximately 21,000 completed interviews, so as to produce reliable estimates for the nation, each state, and DC. Details about this module are included in a separate forthcoming NCHS report and will not be further discussed here.

Sample Design

Like all SLAITS modules, the NSCH took advantage of the large number of screening calls required for the NIS. To accomplish the goal of 1,700 completed NSCH interviews in each state, telephone numbers were initially selected from the telephone numbers randomly generated for the NIS screening effort. Therefore, the procedures for drawing the NIS sample were the first steps in the procedures for drawing the NSCH sample. There were, however, some states for which the NIS sample was not large enough to achieve the desired number of completed NSCH interviews. In these cases, an additional sample was drawn for the purpose of administering the NSCH interview without going through the NIS first. This “augmentation” sample was independent of the NIS, and as a result, it was not subject to NIS screening or interviewing.

The next two sections describe the basic NIS and NSCH sample design, as well as the NSCH sample allocation procedures. [Appendices I, II, and III](#) provide a more technical description of the NSCH sample design and weighting procedures. For more detail on the NIS sample design, readers are encouraged to refer to chapter 2 of the *2007 Methodology Report* for the National Immunization Survey (10), which is available from NCHS. Further information regarding the NIS itself can be found in *National Immunization Survey: The Methodology of a Vaccination Surveillance System* (11) and online from <http://www.cdc.gov/nis>.

The National Immunization Survey Sampling Plan

The NIS monitors vaccination levels of very young children within states and local areas. These “estimation areas” are non-overlapping and cover the United States. Each estimation area is within the borders of a single state, and every location in the US is in one, and only one, estimation area. In effect, each quarter-year, the NIS conducts a separate survey in each estimation area, using a common list-assisted RDD sample design (12, 13). The target number of completed interviews in each estimation area reflects the goal of obtaining equally precise estimates in each estimation area. Thus, the national target for the total number of completed NIS interviews is the sum of the target number of completed interviews in each estimation area. If necessary, the target for an estimation area in a quarter is adjusted to compensate for its total shortfall or excess in previous quarters.

The target population for the NIS is children aged 19 to 35 months, the primary targets of immunization programs. Because less than 5% of households in the US contain children in this age range, the NIS screens over 1 million households per year to identify a sufficient number of households with eligible children. SLAITS modules use this NIS screening sample.

The NIS uses the list-assisted RDD method (12). This method selects a random sample of telephone numbers from “banks” of 100 consecutive telephone numbers (e.g., 673-256-0000 to 673-256-0099) that contain at least one directory-listed landline residential telephone number. The sampling frame of telephone numbers is updated each quarter to reflect new telephone exchanges and area codes.

Although the number of cellular telephone users in the US has increased rapidly, most households with children continue to maintain landline telephone service (14). Also, most cellular telephone users pay for incoming calls. Therefore, the NIS sampling frame excluded cellular telephone banks in 2007.

National Survey of Children’s Health Sample Design and Allocation

The goal of the NSCH sample design was to generate samples representative of populations of children within each state. An additional goal of the NSCH was to obtain state-specific sample sizes that were sufficiently large to permit reasonably precise estimates of the health characteristics of children in each state.

To achieve these goals, state samples were designed to obtain a minimum of 1,700 completed interviews. The number of children to be selected in each NIS estimation area was determined by allocating the total of 1,700 children in the state to each NIS estimation area within the state in proportion to the total estimated number of households with children in the NIS estimation area. Given this allocation, the number of households that needed to be screened in each NIS estimation area was

calculated using the expected proportion of households with children under 18 years of age in the area. Then, the number of telephone numbers that needed to be called was computed using the expected working residential number rate, adjusted for expected nonresponse.

Drawing the Sample

After the number of telephone numbers necessary to achieve the target number of NSCH interviews in each area had been estimated, the samples were drawn. The sample draw proceeded in three steps. First, telephone numbers were sampled for the NIS in each area as previously described. Second, a portion of the telephone numbers in each area was flagged to be part of a supplemental NIS sample designed to assess the vaccination coverage of for teenagers. (This supplemental NIS survey of vaccination rates among teenagers began with a national sample in the final calendar quarter of 2007—Q4/2007—and continued with a state-based sample in 2008.) Third, a portion of the telephone numbers in each area was flagged to be part of the NSCH sample.

After these three steps, every telephone number sampled for the NIS fell into one of four categories: (1) NIS-only, (2) NIS and NIS-Teen, (3) NIS and NSCH, or (4) NIS, NIS-Teen, and NSCH. Households in the fourth group may have been eligible for all three surveys, depending on the ages of the children in the household. To minimize the number of households that would be asked to participate in all three surveys, an effort was made to flag as few NIS telephone numbers as possible for both NIS-Teen and NSCH. Still, it was necessary to have some overlap between the two interviews in the first quarter of 2008 (Q1/2008).

In nine states (Connecticut, Delaware, Idaho, Kansas, Mississippi, Montana, North Dakota, Oklahoma, Utah), there was insufficient NIS sample available to obtain the desired number of NSCH completed interviews. Additional telephone numbers were drawn and categorized as NSCH-only sample. [Table A](#) shows by state the proportion of the overall NSCH sample that was augmented for each state.

Conducting the National Survey of Children's Health Interviews

With the exception of the NSCH-only augmentation sample, each telephone number selected for the NSCH was called and screened for residential status and presence of NIS age-eligible children (including NIS-Teen, if appropriate). NIS interviews were conducted if NIS age-eligible children lived in the household. If NIS age-eligible children did not live in the household, interviewers asked if there were any children younger than 18 years of age who lived in the household. If only one child lived in the household, that child was the subject of the detailed NSCH interview. If there were multiple children in the household, one was randomly selected (i.e., sampled) to be the subject of the detailed NSCH interview.

The sampling and interviewing process described above applied to the entire data collection period except for the first quarter of 2007 (Q1/2007). Households that screened as age-ineligible for the NIS, and households that screened as age-eligible and completed the NIS interview, then moved directly – on the same call, where feasible – to the NSCH age screening and interview. However, in Q1/2007, the NSCH questionnaire was not finalized at the time the NIS was being fielded. Households contacted in Q1/2007 could not move on to NSCH screening and interviewing directly upon finishing the NIS.

As in other quarters, in Q1/2007 a portion of the NIS sample was flagged for the NSCH. After the telephone numbers were called and after the NIS had finished its interviewing efforts, a subsample of this initially-selected sample was selected. This subsample was later dialed for NSCH interviewing. This sample is referred to as the “banked” sample. The subsampling scheme can be found in [Appendix II](#).

Questionnaire

The framework for the 2003 NSCH was initially discussed in September 2001. A panel consisting of state and federal maternal and child health program directors, representatives of family organizations, child health services researchers, and survey design experts met to discuss content domains. Eight domains were selected for their epidemiological and policy importance, including 1) demographics; 2) physical and mental health status; 3) health insurance; 4) health care utilization and access to health care; 5) medical home; 6) family functioning; 7) parents’ health; and 8) neighborhood characteristics. In addition, age-specific modules were identified to capture the developmentally appropriate aspects of child health and well-being. A subset of this panel then assembled questions to capture these domains. Questionnaire items identified for inclusion were then assessed through reviews by outside experts and selected members of the community of potential data users. Upon final approval by MCHB, these questions were pretested in 2002 and fielded in 2003 as the first NSCH.

The 2003 questionnaire underwent revisions prior to implementation in 2007. Revisions to the questionnaire were initially proposed in December 2005 by the Advisory Committee for the Data Resource Center on Child and Adolescent Health and in February 2006 by NSCH data users in response to a request for input distributed by e-mail to members of the SLAITS listserv. Additional input was received in early 2006 from researchers at CDC and at Child Trends. Beginning in March 2006, a technical expert panel reviewed each suggested revision, assembled questions to address newly proposed content areas, and provided recommendations to MCHB. (See [Table B](#) for a list of panel members.) New and significantly revised questions were pretested in December 2006, and the questionnaire was finalized by MCHB and NCHS shortly thereafter.

Content

The NSCH questionnaire was designed to immediately follow a completed NIS interview in households with an NIS-eligible child or the NIS screener in households without an NIS-eligible child. The questionnaire was divided into 11 sections, summarized below.

1. Age-Eligibility Screening and Demographic Characteristics—This section consisted of the introduction to the interview and a question to determine if any children under the age of 18 years were living in the household. All children living in the household were rostered by age, and one child was randomly sampled for the detailed NSCH interview. In this section, respondents were asked questions about their relationship to the sampled child, the sex of the sampled child, and the primary language spoken in the household.

2. Health and Functional Status—The questions in this section were asked to determine whether the sampled child had acute or chronic physical, mental, behavioral, learning, or developmental conditions and, when present, the impact of these conditions upon the child’s life. Respondents were asked

additional questions to determine the presence of various acute and chronic health conditions. This section included the CSHCN Screener, a screening tool developed by the Child and Adolescent Health Measurement Initiative to identify special health care needs in children (15). The CSHCN Screener includes five stem questions on health care needs that could be the consequences of chronic health conditions. If a child currently experiences at least one of those consequences, follow-up questions determine whether each health care need is the result of a medical, behavioral, or other health condition that has lasted or is expected to last for 12 months or longer. Those with affirmative answers to the stem and the follow-up questions are considered to have special health care needs. This screener was also used for the National Survey of CSHCN (5, 8). For identified CSHCN, the NSCH also asked whether the children's condition(s) limit their participation in activities.

3. Health Insurance Coverage—This section established whether the sampled child had adequate health care coverage, and whether there were any gaps in health care coverage during the 12 months prior to the interview. Adequacy was defined by whether the insurance (public or private) meets the child's needs at reasonable costs.

4. Health Care Access and Utilization—Topics in this section included the availability of a usual place for care and a personal doctor or nurse, as well as the need for and use of medical, dental, and mental health services within the 12 months prior to the interview.

5. Medical Home—The main goal of this section was to assess the quality of care from, and communication with, the child's doctors and other health care providers. Together, the items in this section permit an assessment of whether children have access to a "medical home," which is defined by the American Academy of Pediatrics as primary care that is accessible, continuous, comprehensive, family-centered, coordinated, compassionate, and culturally effective (16).

6. Early Childhood (0–5 years)—This section, administered if the sampled child was 5 years old or younger, included questions about developmental screening, child care arrangements, injuries, breastfeeding, use of formula, and solid food introduction. In addition, this section asked about reading, telling stories, watching television, and play. This section included copyrighted questions from the Parent's Evaluation of Developmental Status (PEDS) Child Development Screening Test. The PEDS is a tool to identify children at risk for developmental, behavioral, or social delays (17). Therefore, it was used in this section as a risk assessment tool to identify children who either have or are more likely to have problems. Researchers interested in analyzing the PEDS data should consult the PEDS documentation for scoring instructions. (Health care providers wishing to use PEDS in practice to assess risk status or to make decisions about developmental status for individual children must use the clinical version of the test, which can be obtained from Ellsworth & Vandermeer Press, LLC. The clinical version was not used for the NSCH.)

7. Middle Childhood and Adolescence (6–17 years)—This section, administered if the sampled child was aged 6 years or older, focused on school enrollment, school engagement, activities outside school, social behaviors, and emotional difficulties. Respondents were also asked about their attendance at the sampled child's events and activities; whether they had met all, some, or none of the sampled child's friends; and the amount of time the sampled child spent caring for himself or herself. Questions also

asked about reading, computing, and television watching. This section included the NSCH Social Competence Scale, an internally valid and reliable measure of social skills and behavior problems (18).

8. Family Functioning—This section determined the frequency of family meals and attendance at religious services, as well as the level of stress on the family from the demands of parenting. Three parental stress questions (K8Q31, K8Q32, K8Q34) comprised the Aggravation in Parenting Scale, which was derived from the Parental Stress Index (19) and the Parental Attitudes about Childrearing Scale (20). It has been used previously in the Panel Survey of Income Dynamics, the Survey of Income and Program Participation, and the Survey of Program Dynamics.

9. Parental Health—Questions in this section were designed to obtain the number and type of parents (or people acting as parents) who live with the child and to assess their physical, mental, and emotional health, exercise frequency, and smoking behavior.

10. Neighborhood Characteristics—The primary goal of this section was to ascertain the respondents' perceptions of their neighborhoods, reports about neighborhood amenities and conditions, and beliefs about their children's safety in the neighborhood and in school. Four of the questions in this section (K10Q30, K10Q31, K10Q32, K10Q34) consider parents' perceived level of neighborhood social capital, focusing specifically on positive aspects of social capital relating to children (21). This concept, alternatively called "social support," is similar to the concept of "social cohesion and trust," which is related to variations in violence among inner-city neighborhoods (22). These questions were originally developed for the Longitudinal Studies of Child Abuse and Neglect and have also been used for the Survey of Income and Program Participation.

11. Additional Demographic Characteristics—In this section, respondents were asked a series of demographic questions, including the number of times the family had moved since the child was born, household utilization of assistance from county welfare programs, and the household's ZIP Code. Additional questions determined the race and ethnicity of the child and whether the child and his or her parents were born in the United States. This section also included questions on family income. The annual family income was mapped to Department of Health and Human Services (HHS) Federal Poverty Guidelines for households. This mapping made it possible to categorize the household's income relative to the federal poverty level.

A copy of the questionnaire appears in [Appendix IV](#). [Appendix V](#) lists the key differences between the 2003 and 2007 questionnaires. [Appendix VI](#) lists changes made to the 2007 questionnaire during the data collection period. [Appendix VII](#) includes the DHHS Federal Poverty Guidelines tables used to determine household poverty status and a description of the process to assign poverty status to households. [Appendix VIII](#) contains state-specific health insurance program names used for health insurance questions in Section 3 of the questionnaire.

Programming the Computer-Assisted Telephone Interview

The NSCH was conducted using a CATI system. This software presents the questionnaire on a computer screen to each interviewer and guides the interviewer through the questionnaire, automatically routing the interviewer to appropriate questions based on previous answers. Interviewers enter survey responses directly into the computer; the CATI program determines whether the selected response is

within an allowable range and saves the responses in a survey data file. Online help screens and text are available to aid interviewers. This reduces the time required to transfer, process, and release data, and promotes data accuracy.

The NSCH questionnaire was programmed as a module of the NIS to integrate the two surveys into a single interview. The instrument made full use of the CATI system's ability to check whether a response was within a legitimate range, follow skip patterns, fill state-specific information in questions as applicable (for example, names of state health insurance programs), and employ "pick lists" for response categories. Certain household and demographic questions were identical in both surveys. If a respondent answered these questions during the NIS interview, these questions were not repeated in the NSCH. Instead, answers to these NIS questions were copied to the NSCH data file, as appropriate.

Pretesting

Once initial programming was completed, the instrument underwent rigorous testing to ensure the CATI system functioned correctly. In addition, a pretest of the CATI instrument was fielded in December 2006 with a national list sample of households likely to include children purchased from Survey Sampling International (SSI). Using this list, a total of 640 completed interviews were obtained over the course of 3 days. The administration time for the interview averaged 32 minutes and 37 seconds, with 40 interviews (6%) that lasted longer than 45 minutes. The targeted administration time was 25 minutes (not including the age screening and informed consent process). Based on results from the pretest, changes were made and the questionnaire was finalized.

Interviewer Training

NORC and its subcontractor conducted all interviews for the 2007 NSCH. Interviewer training was conducted by NORC staff at production centers located in Chicago, IL and Las Vegas, NV. The use of multiple sites ensured continuous coverage in all time zones across the US. The number of interviewers who completed training and certification in each month by location is shown in [Table C](#).

The interviewer training sessions for both the pretest and main survey began with an introduction and project overview. Interviewers were informed about project goals, the study purpose and history, sponsors, and design. An overview of each questionnaire section was taught, with emphasis on quality data collection. The relationship between the NSCH and NIS was also covered.

Several exercises on gaining cooperation were conducted throughout training to ensure that interviewers were equipped to answer frequently asked questions (FAQs) and avert refusals. Part of the exercises included pronunciation of medical conditions, as well as a review of the FAQs and other job aids provided for interviewers.

Two types of mock interviews were administered: trainer-led and dual-trainee interviews. The trainer-led mock interviews focused on gaining cooperation skills and the interviewer's project knowledge. The first dual-trainee mock interview was integrated into the section-by-section lecture that progressed through the questionnaire. The interviewers first listened to a lecture regarding each section, and then practiced moving through that section in CATI before discussing the next section. This method

ensured that interviewers became acclimated to the questionnaire, could navigate CATI, and gain cooperation as new topics were introduced. Additional mock interviews were then conducted that simulated more realistic interviewing situations in real time. Each mock interview was designed to highlight various sections of the screener and the main questionnaire, and to provide different scenarios for gaining cooperation.

At the conclusion of the training session, there was an opportunity for trainees to ask any final questions. The class then participated in an interactive game that emphasized project knowledge and FAQs, CATI skills, and common interview situations.

After training, interviewers completed a certification mock interview. This was administered by trained supervisors. It was approximately 30 minutes in length and standardized to ensure that all interviewers were assessed equally in project knowledge, ability to precisely read and correctly pronounce questionnaire items, and ability to answer respondent questions.

A written evaluation was also administered to reinforce what was learned during the training sessions. It was 16 questions in length and took 20 minutes to complete. The evaluation covered FAQs, survey procedures, and question-specific information. Interviewers had to pass both the written and certification mock interview to be certified to work NSCH cases.

Data Collection

Telephone interviewing began on April 5, 2007, and was completed on July 27, 2008. Interviews were completed for 90,557 children and partially completed for an additional 1,085 children. The interview was considered to be at least partially complete if Section 6: Early Childhood (for selected children aged 0-5 years) or Section 7: Middle Childhood and Adolescence (for selected children aged 6-17 years) was completed. See [Table D](#) for the total number of interviews completed and partially completed in each state.

Because 79% of the 91,642 total interviews had been completed by the end of 2007, this survey is referred to as the 2007 NSCH. [Table E](#) shows the total number of interviews completed by month.

Advance Letters

Advance letters have been shown to decrease nonresponse; they confirm study legitimacy and communicate the value of the survey (23). When a mailing address could be identified for a sampled telephone number, an advance letter was mailed prior to any telephone calls. Letters were mailed for 58.8% of the telephone numbers eventually dialed by the interviewers, which was 31.0% of the total telephone numbers randomly generated. (Some known business and nonworking telephone numbers are removed from the sample of randomly generated telephone numbers prior to dialing.) [Appendix IX](#) contains the full complement of advance letters used over the course of data collection.

Because the NSCH typically follows the NIS, the advance letter sent to most households was the usual NIS advance letter. It asked recipients to participate in a voluntary study on the immunization status of their children and the types of health and related services that their children need and use. The letter also explained how their telephone number was selected, who was conducting the survey, and that

their household would be contacted within the next two weeks. The letter included an additional page of frequently asked questions covering topics such as confidentiality and the legitimacy of the survey. The letter provided toll-free telephone numbers for those with concerns or questions about the study and for respondents who wanted to participate immediately. Finally, the letter offered a website address for more information about the NIS.

As described earlier, the NIS sample was augmented with additional sample in states where NIS sample was insufficient to complete the required number of NSCH interviews. These households were sent an advance letter specific to the NSCH. The advance letter for this augmentation sample explained that the interview would include questions on health topics such as exercise, sleep, diet, and doctor visits for children and teenagers, and that this information would be used to help create programs for healthier children, schools, and communities. As with the NIS advance letter, recipients were advised that their telephone numbers had been chosen randomly and they might be called in the next few weeks. Households in the augmentation sample were given a unique NSCH-only toll-free number to call if they wished to participate immediately or to learn more about the study. The letter did not mention anything about the NIS or immunizations, and it gave the address for the NCHS SLAITS website.

Toll-Free Telephone Numbers

A toll-free telephone line offered respondents the flexibility to call at their convenience if they had questions about the survey or wanted to establish eligibility, complete the interview, or submit feedback. Advance letters, incentive letters, answering machine scripts, and closing scripts referenced this toll-free number, and interviewers provided the number to respondents who requested it during the interview.

The telephone line was answered by NSCH interviewers. During the course of data collection, 8,822 cases in the NSCH sample called the toll-free line. Out of these cases, 4,918 households were determined to be ineligible, and an additional 2,710 households were screened as eligible. Overall, respondents in 2,395 of these eligible households ultimately completed the NSCH.

A second toll-free telephone number connected recipients of the letter to the NCHS Research Ethics Review Board (ERB) for answers to questions about survey legitimacy, confidentiality, and the rights of respondents. The ERB number was provided in the advance letters and in the closing interview script in case respondents had questions after completion.

Selection of Sampled Child

Households were screened for the presence of children less than 18 years of age. In households with children, the ages of all children living or staying in the household were then obtained. If a household only had one child, that child was selected as the focus of the interview by default. In households with multiple children, one child was randomly selected to be the focus of the interview.

Selection of Respondent

Interviewers requested to speak with a parent or guardian who lived in the household and who knew about the health and health care of the sampled child. The respondent was the mother or father for 94% of sampled children. [Table F](#) shows the frequency of respondents by their relationship with the sampled child.

An adult 18 years of age or older was not identified in 2,843 households (0.5% of all known households contacted for the NSCH). These households were not eligible to complete the screening portion of the interview. No interviews were conducted in these households even if a minor who lived there was the parent of a younger child.

Informed Consent

After the knowledgeable adult came to the telephone, or after the person who answered the telephone identified herself or himself as the knowledgeable adult, this respondent was informed of her or his rights as a survey participant. Verbal consent for study participation was then obtained and documented in the CATI system. The informed consent script told respondents of the voluntary nature of the survey, assured them their responses would be kept confidential, indicated there was no penalty for not answering questions, and informed them about survey content and expected duration. If the respondent was eligible for a monetary incentive payment (see “Incentive Effort” below), they were also told they would receive \$10 or \$15 in appreciation of their time. Finally, the respondent was also told the interview might be recorded and monitored by a supervisor for quality control.

In accordance with DHHS regulations (45 CFR 46), the NCHS ERB and the NORC Institutional Review Board (IRB) approved all study procedures and modifications. The Federal Office of Management and Budget control number for this collection of information was 0920-0406.

Assurance of Confidentiality

Participation in surveys conducted by NCHS is voluntary, and all data collected that could potentially identify an individual person are confidential. For the NSCH, assurance of confidentiality was provided to potential respondents as part of the informed consent procedures. In the CATI system, interviewers acknowledged that they read the following statement to respondents:

Before we continue, I'd like you to know that taking part in this research is voluntary. You may choose not to answer any questions you don't wish to answer, or end the interview at any time. We are required by Federal law to develop and follow strict procedures to protect your information and use your answers only for statistical research. I can describe these laws if you wish. [In appreciation for your time, we will send you \$10/\$15.] The survey will take about 25 minutes. In order to review my work, my supervisor may record and listen as I ask the questions. I'd like to continue now unless you have any questions.

If respondents requested to hear more about these Federal laws, they were read the following statements:

The Public Health Service Act is Volume 42 of the US Code, Section 242k. The collection of information in this survey is authorized by Section 306 of this Act. The confidentiality of your responses is assured by Section 308d of this Act and by the Confidential Information Protection

and Statistical Efficiency Act. Would you like me to read the Confidential Information Protection provisions to you?

If the respondent indicated that he or she would like to hear the Confidential Information Protection provisions, the interviewer read the following statement:

The information you provide will be used for statistical purposes only. In accordance with the Confidential Information Protection provisions of Title V, Subtitle A, Public Law 107-347 and other applicable Federal laws, your responses will be kept confidential and will not be disclosed in identifiable form to anyone other than employees or agents. By law, every employee of the National Center for Health Statistics, the National Center for Immunization and Respiratory Diseases, and its agent, the National Opinion Research Center who works on this survey has taken an oath and is subject to a jail term of up to 5 years, a fine of up to \$250,000, or both, if he or she willingly discloses ANY identifiable information about you or your household members.

When NCHS (including its contractors and agents) collects personally identifiable information under a pledge of confidentiality for exclusively statistical purposes, Section 308d of the Public Health Service Act and Section 512b of the Confidential Information Protection and Statistical Efficiency Act (CIPSEA) require without exception that the confidentiality of the personally identifiable information be maintained throughout the research and thereafter. Strict procedures are used by NCHS, its data collection contractors, and other agents to prevent any disclosure of confidential data in survey operations and data dissemination.

Identification of Age-Eligible Children for the National Immunization Survey

On occasion, a household indicated that there were no NIS-eligible children in the household, but upon rostering the children's ages in the NSCH, potentially NIS-eligible children were found. When this occurred, the interview returned to attempt to complete the NIS first, prior to continuing with the NSCH interview. There were 201 such households identified in the NSCH, and through re-screening in NIS, 146 of these households were determined to be age-eligible for the NIS. Of the 201 cases, 122 cases then returned to the NSCH interview after the NIS screening or the NIS interview was completed. Of these 122 cases that returned to the NSCH, 82 completed the NSCH interview.

Spanish-Language Interviewing

NSCH interviews were administered in Spanish as well as in English. A professional translator produced a Spanish-language version of the NSCH questionnaire, and an independent translator checked the accuracy of the original translation. A team of experienced Spanish-speaking telephone interviewers and supervisors at NORC reviewed the translation and evaluated it for accuracy and cultural appropriateness. Issues raised during this review were resolved in consultation with the original translator. Any necessary modifications were made, and the translated questionnaire was programmed into the CATI system for testing and eventual production.

All households were first called by an English-speaking interviewer. If a respondent answered the telephone in a language other than English, interviewers asked, "What language do you speak?" If it was determined that the respondent needed a Spanish-speaking interviewer, the case was placed in a Spanish

calling queue. If the interviewer placing the initial call was a Spanish speaker and trained to administer the Spanish version of the questionnaire, the interviewer toggled to the Spanish questionnaire and continued the interview without interruption. If not, the call was terminated, the case was flagged in the CATI system as needing a Spanish interviewer, and all subsequent calls were made by Spanish-speaking interviewers. Nevertheless, the interview may have been conducted in English if a subsequent call by a Spanish interviewer reached an English-speaking respondent.

During data collection, 21,853 telephone numbers were placed in the Spanish calling queue. Of these, 18,067 were determined to reach households, and 14,813 of these households were screened for age-eligibility. Of 7,923 households with age-eligible children, 4,828 completed the NSCH interview (4,672 full completes and 156 partial completes). Households placed in the Spanish queue comprised 5.9% of all screened households with children and 5.3% of all completed NSCH interviews. At the end of the NSCH interview, the interviewer recorded the language used to conduct the interview. Of the 4,672 full completes from cases placed in the Spanish queue, 4,407 completed the interview in Spanish.

Interviewing in Languages Other than Spanish or English

Based on the experience of the 2005-2006 National Survey of CSHCN, four languages were identified as the most probable languages that interviewers would encounter other than English or Spanish: Mandarin, Cantonese, Vietnamese, and Korean. Independent translators translated the NSCH questionnaire into these Asian languages using the same procedures as were used for the Spanish questionnaire. Although the Spanish questionnaire was programmed into the CATI system, given the expected low incidence of the other languages, a different procedure was followed to screen and interview these Asian language households.

When a household was first identified as needing a language other than English or Spanish, the case was transferred to specially trained interviewers who would determine the necessary language with a language service used by NORC, Language Line Services. Language Line Services provides a real time translation service in more than 170 languages. These households were then screened for NIS age-eligible children, and if they were eligible for the NIS, the interviewer immediately conducted the NIS interview with the assistance of the Language Line interpreter. After a completed NIS interview, or after NIS age screening if there were no NIS age-eligible children living in the household, the interviewer (with the help of the interpreter) screened the household for children under 18 years old. If the household included children and spoke one of the four targeted Asian languages, the case was assigned to the appropriate language queue to be called by a specially trained interviewer who spoke that language. Special language interviewers entered the respondent's answers into the regular English CATI system, while using a book that contained the translated questionnaire. This allowed for the data to be captured immediately in the CATI system and to be subject to all computerized logic and validation checks.

Throughout the course of data collection, 416 households were identified as needing an interview in one of the four available Asian languages. Of these, 12 were determined to be age-ineligible and 394 were determined to be age-eligible. The NSCH interview was completed with 167 of the age-eligible households (156 full completes and 11 partial completes). Households that were identified as needing an Asian language interviewer comprised 0.3% of all screened households with children, and 0.2% of all completed NSCH interviews. At the end of the NSCH interview, the interviewer recorded the languages

used to conduct the interview. Of the 156 full completes from cases identified as needing an Asian language interview, 123 completed the interview in one of the Asian languages.

If the Language Line interpreter reported that a household included age-eligible children but did not speak English, Spanish or one of these four Asian languages, the case was coded as “age eligible, interview incomplete” and the case was finalized. A total of 688 households with children were finalized due to language.

Interview Length

Mean and median interview length varied by NIS or NIS-Teen eligibility because some demographic and household questions necessary for the NIS, NIS-Teen, and the NSCH were administered as part of the NIS or NIS-Teen interview and not repeated during the NSCH interview. The average interview length for NIS-ineligible and NIS-Teen-ineligible households was 28 minutes, 5 seconds, and the median time was 26 minutes, 56 seconds. For NIS-eligible or NIS-Teen-eligible households, the average interview length (excluding the NIS and NIS-Teen interviews) was 22 minutes, 55 seconds, and the median time was 21 minutes, 54 seconds. Mean and median interview lengths, by section and NIS and NIS-Teen eligibility, appear in [Table G](#).

Interview Breakoffs

Households that terminated the interview before completion were placed into a queue that was worked by interviewers specially trained in refusal conversion strategies. These interviewers attempted to convert the incomplete interview into a completed interview. By the end of data collection, 12,619 interviews were completed with households that had refused to participate at least once after age-eligibility was established (13.8% of completed interviews).

There were 43,734 identified households with children in which an interview was not completed (1.6% of the initial sample and 32.3% of identified households with children). Of these households, 3,138 broke off during the NIS interview and 840 broke off during the NIS-Teen interview. The remaining 39,756 cases reached the NSCH interview. The most common places in the NSCH interview where respondents broke off were during the screener at the rostering of children’s ages and the sampling of a child (22.6% of such cases), during the question asking for a respondent knowledgeable about the health and health care of the sampled child (21.0%), or during the informed consent script (32.9%). Among the 9,336 cases that prematurely ended the interview after consenting to continue during the informed consent process, 4,649 cases broke off before the first question in Section 1 (11.7% of the breakoff cases). For the 4,687 cases (11.8% of the breakoff cases) that began the interview, there was little commonality in the location of the questionnaire where the interview was terminated.

Cases Pending at Close of Data Collection

Most of the cases pending at the end of data collection were those in which the telephone number had not yet been resolved as residential or nonresidential (79.1% of pending cases and 15.9% of the initial sample). A smaller number of cases had been resolved as households without age eligibility being determined (13.1% of pending cases and 2.6% of the initial sample). Finally, as noted previously, 32.3% of all age-eligible households had not completed the interview (7.8% of pending cases and 1.6% of the initial sample). See [Table H](#) and [Appendix X](#) for more information about final dispositions of cases.

Incentive Effort

In a proactive effort to address nonresponse by NSCH-eligible households, the NSCH implemented a cash incentive plan shortly after data collection began in April 2007. A two-tiered incentive model was constructed based on the productive incentive efforts executed in the 2003 NSCH and the 2005-2006 National Survey of CSHCN.

Careful monitoring of key response rates during the first quarter of data collection identified a group of cases that were either not responsive to the initial incentive offer or had characteristics that precluded them from incentive eligibility. An expanded incentive effort began in August 2007 to reach out differentially to these households with an incentive offer.

Interview completion rates were favorable for the two-tiered model and the expanded incentive model, with particularly high completion rates among the expanded incentive-eligible cases in the final two quarters of data collection. More detailed descriptions of the incentive models, the process by which cases were offered an incentive, and the completion rates are included in [Appendix XI](#).

Response Rates

Response rates provide one measure of the potential for nonresponse bias – that is, the possibility that the sample interviewed differs from the target population in some meaningful way. The NSCH weighted response rate, calculated nationally and by state, reflects the potential for bias in the sample of children for whom the interview was completed.

The response rate was calculated in accordance with the American Association for Public Opinion Research's (AAPOR) standards for Response Rate #4 (24). This response rate calculation recognizes that some cases of unknown eligibility (e.g., telephone lines that rang with no answer, or households in which the person answering the phone refused to say whether the household included children) were in fact eligible. In accordance with Council of American Survey Research Organizations (CASRO) guidelines, the proportion of eligible cases among those with unknown eligibility was assumed to be the same as the proportion of eligible cases among those with known eligibility. The response rate was calculated as the product of component completion rates, which are discussed below.

We also provide an alternative set of completion rates and overall response rates. The alternative response rates were calculated similarly, with the exception that, in the calculation of the alternative rates, it was assumed that cases that were never contacted are not households.

Response rates are just one measure of the potential for nonresponse bias. A more formal examination of nonresponse bias is underway and will be published by NCHS later this year.

Resolution rate

Response rates for household telephone surveys are typically lower than response rates for household in-person surveys because some telephone numbers ring with no indication of whether the number belongs to a household or to a business. The NSCH's national resolution rate, which measures the proportion of sampled telephone numbers that could be identified as residential or non-residential, was 81.9%. Resolution rates for each state, which ranged from 77.2% to 89.2%, are given in [Table J](#). When called, the majority of the unresolved telephone numbers rang with no answer. Most of the other

unresolved numbers either reached persons or machines who “hung up” before identifying themselves or reached answering machines that provided no indication of whether the caller had reached a residence or a business.

Screener completion rate

After a telephone number had been determined to belong to a household, that household was screened for the presence of children under 18 years of age. Each household (except for those in the augmentation sample) was first screened for NIS eligibility; that is, each household was screened for the presence of children age 19 to 35 months (and, for some Q1/2008 sample, for children 13 to 17 years). If a household was age-eligible for the NIS, then the household was also considered to be age-eligible for the NSCH. If a household was age-ineligible for the NIS, then that household may or may not have been age-eligible for the NSCH and so proceeded to the NSCH age screening. If the respondent then indicated that the household contained children under age 18 (at question S_UNDR18), the household was considered to be age-eligible for the NSCH. If, during the NIS or NSCH age screeners, the household indicated that they had no children whatsoever, the household was considered to be age-ineligible for the NSCH.

For some households, it was never determined whether the household contained children under age 18 years; that is, some households did not complete the NSCH age-screener. The screener completion rate is defined as the proportion of identified households for which it was determined whether or not the household contained children under age 18 years. The national screener completion rate was 86.4%. Screener completion rates for each state, which ranged from 83.4% to 90.6%, are listed in [Table J](#).

Interview completion rate

After it had been determined that a household contained children under 18 years of age, a child was randomly chosen from the household, and an attempt was made to conduct a full interview about the selected child. As noted previously, the interview was considered to be at least partially complete if Section 6: Early Childhood (for selected children aged 0-5 years) or Section 7: Middle Childhood and Adolescence (for selected children aged 6-17 years) was completed.

Not all households containing a child completed the NSCH interview. The interview completion rate is defined as the proportion of age-eligible households that completed Section 6 or Section 7 of the NSCH interview. The national interview completion rate was 66.0%. Interview completion rates for each state, which ranged from 60.5% to 76.6%, are listed in [Table J](#).

Overall Response Rate

The overall response rate is the product of the resolution rate, the screener completion rate, and the interview completion rate. At the national level, the response rate was 46.7%. Overall response rates for each state, which ranged from 39.4% to 61.9%, are listed in [Table J](#).

Alternative Response Rates

The overall response rate just presented was based on a conservative approach to estimating the proportion of age-eligible households among the cases (i.e., telephone numbers) of unknown eligibility. This approach assumed the same working residential rate for the unresolved telephone numbers as was

observed among the resolved telephone numbers, and it assumed the same age-eligibility rate for the non-age-screened households as was observed among the age-screened households.

An alternative response rate that still meets the general criteria for AAPOR Response Rate #4 can be calculated by changing the assumption that the working residential rate for the unresolved telephone numbers was the same as was observed among the resolved telephone numbers. A less conservative assumption holds that all telephone numbers that resulted in no contact (i.e., all attempts resulted in rings with no answer or in a busy signal) were not working residential numbers. Because every telephone number is dialed at least six times at different times on different days, the assumption that all of these “noncontact” numbers are not working residential numbers is somewhat tenable. A portion of all other unresolved cases are considered eligible, using the same working residential number rate as was observed among the resolved telephone numbers.

When this alternative assumption was applied nationally, the alternative resolution rate was 89.9%, which is necessarily higher than the 81.9% resolution rate in [Table J](#) because noncontact cases were assumed to be nonresidential. The component screener completion and interview completion rates remain the same, at 86.4% and 66.0%, respectively, because noncontact cases do not enter into the calculation of these rates. As a result, the alternative national overall response rate of 51.2% is higher than the 46.7% overall response rate. The alternative rates for each state, which ranged from 44.9% to 64.9%, are given in [Table K](#).

Researchers choosing to report these alternative response rates should clearly state the assumption on which these alternative response rates are based. For example, researchers could include the following statement in written reports that use the alternative response rates: “Numbers that, with six or more call attempts on different days and times, rang with no answer or were busy on all attempts were assumed to be nonworking or nonresidential numbers.”

Efforts to Maximize Response Rates

Advance letters, toll-free telephone numbers, cash incentives, refusal conversion efforts, and translated questionnaires were used to help maximize response rates. In addition, a pretest was conducted to understand how respondents would react to potentially personal questions and a lengthy interview, and to monitor respondent suspicions of legitimacy and confidentiality, among other issues. After the pretest results were analyzed, specific improvements were made based on these findings. After every quarter of data collection, NORC and NCHS reviewed and (if necessary) implemented interviewers’ and supervisors’ recommendations for potential changes to the questionnaire, data collection procedures, and calling rules. These changes were based on analysis of questionnaire breakoffs and reports from interviewers of problem areas within the questionnaire.

Quality Control of Interviewing

Telephone center supervisors were available to interviewing staff at all times to resolve any questions or concerns about a case. Supervisors regularly observed the data collection process to monitor interviewers informally. In addition, supervisory staff used remote telephone and computer-monitoring technology to evaluate whether interviewers performed according to project specifications. This formal monitoring was conducted to ensure that introductory materials were properly read, that item wording and

sequence of the questionnaire were followed correctly, that respondent questions were answered properly, and that any vague responses were properly probed. Computer monitoring also allowed supervisors to ascertain whether answers were entered accurately into the CATI system.

Supervisory staff monitored approximately 10% of all calls made for the NSCH. To avoid bias in selecting whom to monitor, the CATI monitoring system automatically selected which interviewers to monitor, using an algorithm that gave the highest priorities for selection to newly trained interviewers, those with the fewest monitoring sessions, and those with the weakest performance reviews. Experienced interviewers were prioritized for monitoring based upon the length of time since their last monitoring session and recent monitoring scores. Each interviewer was typically monitored at least once a week;.

Throughout data collection, interviews were recorded (after obtaining agreement from respondents). These recordings were valuable tools for trainings, and when necessary, they allowed supervisors to document specific case-related performance issues.

Data Files

One SAS data file (version 9.1) was created using data from completed and partially completed interviews that were conducted in 2007 and 2008. The interview was considered to be at least partially complete if Section 6: Early Childhood (for selected children aged 0-5 years) or Section 7: Middle Childhood and Adolescence (for selected children aged 6-17 years) was completed. This file contains data on each sampled child's health and health care, health insurance, family functioning, parental health, and neighborhood/community characteristics. There is one record for each child that was randomly selected to be the subject of the interview. Each record contains all interview data for the sampled child and the household in which the child resides. Of the 91,642 records, 90,557 cases completed the full interview, and 1,085 were partially completed interviews.

An additional SAS data file includes multiply imputed household poverty data. Details about the imputed poverty data are included in [Appendix XII](#).

Editing

As discussed previously, the CATI system was designed to perform edits as an interviewer entered data into the computer system. To prevent interviewer error, the CATI system was developed to include range checks and consistency checks. If an interviewer entered a value that was "out of range," a warning screen would appear, instructing the interviewer that the value would not be accepted and that he or she would have to enter a different answer (and possibly re-ask the question). For example, a respondent might report three people living in the household, but if the respondent had earlier reported four children, a consistency check would appear saying that the number of people living in the household must be greater than the number of children.

Despite these range checks and consistency checks, some respondents still provided logically inconsistent responses. Interviewers were trained not to challenge respondents who gave logically inconsistent responses. Logically inconsistent responses given by the respondent were left inconsistent in the data files.

Data cleaning was necessary to delete invalid values and investigate missing values. Most missing values were the result of legitimate skip patterns within the questionnaire or the result of a partially completed interview. On rare occasions, certain data were not collected as expected. If, based on related questions, the missing data were easy to determine, the correct answers were added. Records that were missing responses for unknown reasons were left missing.

Missing Data

Missing data are not desirable when doing analyses, and are often ignored completely. However, it can be very helpful to know why data are missing. The SAS data file for the NSCH includes special missing value codes for analysts who may wish to differentiate between different types of missing values. The following key provides a description of the various codes that were used to represent missing data in the file.

(.N) Not in universe—Respondents skipped an entire section of questions based on eligibility criteria. For the NSCH, sampled children ages 0–5 years were not eligible for Section 7 of the survey, and children ages 6–17 years were not eligible for Section 6 of the survey.

(.L) Legitimate skip—Variable is missing due to valid questionnaire paths based on a previous answer to a root question.

(.P) Partially completed interview—Variable is missing because the respondent ended the interview after completing Section 6 or Section 7 but before completing the full interview.

(.M) Missing in error—Variable is missing due to interviewer or system errors. In cases of interviewer error, the interviewer may have deleted the data by accident or simply may have not entered the response. In cases of system error, the response may not have been collected or saved properly after it was entered by the interviewer in the CATI system.

(.A) Added question—Variable is missing because this question was added after the start of data collection and the interview was conducted before the question was added.

(.D) Deleted question—Variable is missing because this question was removed after the start of data collection and the interview was conducted after the question was deleted.

Because SAS treats all of the above codes similarly in statistical analyses (i.e., as missing data), analysts using SAS who are not interested in the reasons for the missing data may continue to analyze data as usual.

It is important to note that derived variables (i.e., new variables calculated from responses directly provided by the respondent) do not undergo detailed coding for missing data. All missing values for derived variables received “.M” codes regardless of the reason for the missing data. Similarly, “.M” was used when derived variables were suppressed to protect the confidentiality of the survey participants.

Data missing because the respondent did not know the answer or refused to provide the answer have been treated differently. Rather than assigning a missing value to these records, a numeric code was used to identify these responses. Typically, unknown answers are coded as “6,” “96,” or “996.” Refused

responses are coded as “7,” “97,” or “997.” However, the codes may be different for specific variables; therefore, analysts are encouraged to consult the data documentation and frequency lists to identify the correct codes for each variable. Failure to do so may result in inappropriate calculations, especially for variables measured using ordinal, interval, or ratio scales.

Coding of Verbatim Answers into Question Responses

For some questions in the NSCH interview, respondents provided a response that did not match any pre-existing category. If this occurred, the interviewer chose “other” and typed in the exact response provided by the respondent, to the extent possible. At the end of the data collection period, the verbatim responses were recoded as necessary into existing response categories.

This recoding occurred for respondents who did not choose one of the pre-existing categories when reporting race and ethnicity. This recoding also occurred for respondents who did not choose one of the pre-existing categories for the child’s usual place of care when sick (K4Q02R).

Edits to Protect Confidentiality

NCHS takes extraordinary measures to assure that the identity of survey subjects cannot be disclosed. The risk of inadvertent disclosure of confidential information regarding individual respondents is higher with a publicly released data set having detailed geography variables, a detailed and extensive set of survey observations, and a sizeable proportion of the total population of interest. Coarsening a data set by suppressing survey variables, collapsing multiple variables into one, collapsing response categories, and/or introduction of noise in the data are common techniques to reduce the risk of inadvertent disclosure.

In these data files, the child’s exact age (in months) has been suppressed, but the child’s age (in years) has been reported, along with indicators to identify children less than 6 months of age (FLG_06_MNTH) and less than 18 months of age (FLG_18_MNTH). The specific relationship of the respondent to the child (RELATION) has been suppressed when the respondent was not the parent of the child. Household income has been suppressed, but a measure of income relative to the federal poverty level has been included (POVERTY_LEVELR). The date of the interview has been suppressed, but for those researchers concerned about seasonality effects in responses, an indicator (SUMMER) identifies those interviews that were completed when children were more likely to not be in school.

Geography

Geographic information that would identify the specific estimation area in states with multiple estimation areas has been suppressed. However, state identifiers are included. In addition, an indicator identifying whether or not the household resides inside or outside of a Metropolitan Statistical Area (MSA) has been included for some states. This indicator, called MSA_STAT, was suppressed whenever the total population for all MSA areas in a given state was less than 500,000 persons, or whenever the total population for all the non-MSA areas in a given state was less than 500,000 persons. This resulted in the suppression of the MSA identifier in 16 states.

Because the MSA identifier is suppressed in 16 states, national estimates by MSA status are not possible with the publicly available NSCH dataset, and analysts should use caution when including this

variable in statistical models. Analysts may consider using imputation to assign an MSA indicator to children in states where the indicator was suppressed. One option for analyses at the national level is to assign MSA status to children in states that are predominately metropolitan and to assign non-MSA status to children in states that are predominately non-metropolitan. If MSA status is imputed to all children in Alaska, Connecticut, Delaware, Hawaii, Idaho, Maine, Maryland, Massachusetts, New Hampshire, Nevada, and Rhode Island, the MSA identifier will be correct for 79% of the children (15,582 out of 19,684). If non-MSA status is imputed to all children in Montana, North Dakota, South Dakota, Vermont, and Wyoming, the MSA identifier will be correct for 63% of the children (5,509 out of 8,801). For weighted national analyses, this imputation procedure will result in erroneous classifications for 8.1% of children.

Race

Question K11Q02 asked about the sampled child's race. Respondents were permitted to identify all possible categories that described the child's race. If a race other than one of the seven existing categories was indicated, then a verbatim response was captured. Verbatim responses were reviewed and matched against a database of alternative race terminology maintained by the U.S. Census Bureau. Where possible, "other" race responses were backcoded into one of the seven existing categories. Once all possible verbatim responses were backcoded, a new race variable was created by collapsing the seven categories into one of six categories: White, Black or African American, American Indian or Alaska Native, Asian, Native Hawaiian or Other Pacific Islander, and multiple race. "Multiple race" was reserved for those cases where more than one of the other five categories applied.

To protect the confidentiality of individual respondents and children, responses for the race variable were further collapsed into four categories: white only, African American or black only, other race, and multiple race. This "other race" category includes children for whom only one of the other three categories (Asian, American Indian or Alaska Native, and Native Hawaiian or Other Pacific Islander) was reported. Children for whom more than one race was identified (e.g., Asian as well as Native Hawaiian) were included in the "multiple race" category. If the respondent did not know or refused to provide the race, then race was coded as missing. Cases where a verbatim response could not be conclusively backcoded (e.g., American, Indian, Jewish) and no other race was reported were also coded as missing. This new derived race variable (called RACER) is the only classification publicly available for all 50 states and DC.

In several states, however, minority group populations are sufficiently large that the release of additional race categories was possible while still protecting the confidentiality of the respondents and children. To identify these states, data from the 2007 American Community Survey (ACS) were examined to identify minority groups that comprise at least 5% of the total population of children in a specific state. Based on this criterion, the data files identify American Indian or Alaska Native children in Alaska, Arizona, Montana, New Mexico, North Dakota, Oklahoma, and South Dakota. This race classification variable is called RACEAIAN. Asian children's race is reported for children in California, Maryland, Massachusetts, Nevada, New Jersey, New York, Virginia, and Washington. This race classification variable is called RACEASIA. The data file identifies both Asian children and Native Hawaiian or Other Pacific Islander children in Hawaii. This race classification variable is called RACE_HI.

It is important to note that national estimates for American Indian or Alaska Native children, for Asian children, and for Native Hawaiian or Other Pacific Islander children are not possible using the publicly available NSCH dataset. Children with these race classifications are identified in selected states only. These race classifications were suppressed in other states for 583 American Indian or Alaska Native children, 1,245 Asian children, and 258 Native Hawaiian or Other Pacific Islander children. These children with suppressed race classifications represent 62.4% of American Indian or Alaska Native children, 37.1% of Asian children, and 85.9% of Native Hawaiian or Other Pacific Islander children nationally.

Language

Question K1Q03 collected data on the primary language spoken in the household. Of the 6,643 children living in households with a non-English language as the primary language (PLANGUAGE), 76.0% (n = 5,052) lived in Spanish-language households. Of the remaining non-English-language households, 352 (5.3%) spoke one of the four Asian languages in which the interview was conducted and 1,239 (18.7%) spoke another language. To protect confidentiality, the specific language spoken in non-English-language households and the specific language used for any non-English interview (OTH_LANG) have been suppressed.

Height and weight

Question K2Q02 permitted respondents to report the child's height in either feet and inches or in centimeters. Height reported in centimeters was recoded into inches (K2Q02R). Question K2Q03 permitted respondents to report the child's weight in either pounds or kilograms. Weight reported in kilograms was recoded into pounds (K2Q03R). Question K2Q04 permitted respondents to report the child's birth weight in either pounds, ounces, or grams. Weight reported in pounds and grams were recoded into ounces (K2Q04R).

To protect the confidentiality of individual children, very short heights, very tall heights, very low weights, and very high weights have been suppressed. Extreme values were identified within each single-year age group and were recoded to less-extreme values. For example, for 11-year-old children, all reported heights shorter than 43 inches were recoded to 43 inches, and all reported heights taller than 68 inches were recoded to 68 inches. Two flags (HGHT_FLG and WGHT_FLG) have been added to the dataset to enable analysts to determine whether the values were reported or assigned.

Because suppression of height and weight variables may hinder calculations of body mass index (BMI), a 4-category variable identifying underweight and overweight children (BMICLASS) has been added to the dataset. Children aged 10-17 years have been identified as having a BMI-for-age that is equal to the 5th percentile or lower, greater than the 5th percentile but lower than the 85th percentile, equal to the 85th percentile or greater but lower than the 95th percentile, and equal to the 95th percentile or greater. Percentiles are based on gender and age (25). For example, if the value of a child's BMI is equal to the 95th percentile, then that child is among the 95% of children of that age and sex whose BMI is less than or equal to that value. Percentiles were determined using the 2000 CDC growth charts and a SAS statistical analysis software program provided on-line by CDC (26). However, this program relies on the child's age in months; because age was only reported in years for this survey, children were

assumed to be at the midpoint of the age-year (i.e., a 10-year-old was assumed to be 126 months of age) for purposes of calculating BMI-for-age.

Height and weight were based on parent report and were not independently measured. Researchers attempting to validate parent report of height and weight in the 2003 NSCH have concluded that parent-reported data should not be used to estimate overweight prevalence among preschool-aged and elementary school-aged children (27). Parents' reports significantly underestimated height; as a result, too many young children were classified as overweight in the 2003 NSCH. Due to concerns about the validity of the 2007 data, reported height (K2Q02R) and calculated BMI categorizations (BMICLASS) have been suppressed for children less than 10 years of age.

Family structure

To protect the confidentiality of individual children whose families have unique structural characteristics, a single measure of family structure (FAMSTRUCT) was created from K1Q02, K9Q00, K9Q10, K9Q11, and K9Q12. The family structure variable refers to parents living in the household. This variable has four levels: 1) two-parent household with both a biological or adoptive mother and a biological or adoptive father; 2) two-parent household with both a mother and a father that includes at least one step-parent; 3) one-parent household with a biological, step, foster, or adoptive mother and no father of any type present; 4) all other family structures. Any of these four family structures may include other people who act as parents, such as grandparents, aunts, uncles, or unmarried partners of the parents. Legal guardians were not considered to be mothers or fathers.

Households identified as having two mothers of the same type (biological, step, foster, or adoptive) have been classified as "other family structure." Households with an ambiguous structure (e.g., where a father refused to indicate whether he was the biological father) were also coded as "other family structure."

In addition, variables were suppressed that could be used to identify households where the child's biological parents were still married to each other but not living together (K9Q17B_1, K9Q17C_1, K9Q17D_1), as well as the specific reasons for not living together (K9Q17B_2, K9Q17C_2, K9Q17D_2).

Top-coded and bottom-coded variables

Several other frequency variables have been top-coded and/or bottom-coded to suppress outliers at the high and low ends of the distributions of responses. Due to their unusual characteristics, cases including these outliers might have been more readily identifiable. The minimum and maximum categories released on the publicly available data files are listed below.

- For the total number of children living in the household (TOTKIDS4), "4 or more" children is the maximum category released publicly.
- For the total number of adults living in the household (TOTADULT3), "3 or more" adults is the maximum category released publicly.
- For the child's birth weight (K2Q04R), "47 or fewer" ounces is the minimum category and "163 or more" ounces is the maximum category released publicly.

- For the number of visits to a doctor, nurse, or other health care professional for preventative medical care in the past year (K4Q20R), “20 or more” visits is the maximum category released publicly.
- For the number of visits to a dentist for preventative dental care in the past year (K4Q21R), “12 or more” visits is the maximum category released publicly.
- For the age of the child when breastfeeding stopped (K6Q41R), “1095 or more” days (i.e., 3 years or older) is the maximum category released publicly.
- For the number of days of school missed due to illness or injury in the past year (K7Q02R), “40 or more” days is the maximum category released publicly.
- For the number of times that the parent was contacted by the school during the past year (K7Q04R), “25 or more” times is the maximum category released publicly.
- For the frequency that the child attended religious services in the past year (K8Q12R), “daily or more often” is the maximum frequency released publicly.
- For the age of the child’s mother (K9Q16R), “20 or fewer” years in the minimum category and “59 or more” years is the maximum category released publicly.
- For the length of time that the child or parent had been living in the United States (K11Q34AR through K11Q37AR), the variable has been recoded as number of years, and “one year or less” is the minimum length of time released publicly. For parents, “45 years or more” is the maximum length of time released publicly.
- For the number of times that the child ever moved to a new address (K11Q43R), “12 or more” times is the maximum category released publicly.
- For the education level of the mother, father, and respondent (EDUC_MOMR, EDUC_DADR, and EDUC_RESR), post-high school study is the maximum category released publicly, and other responses have been collapsed into two additional categories (less than high school graduate, high school graduate or GED completed).

Data perturbations

Despite the modifications detailed above, there was lingering concern that the dataset may include children with unique combinations of identifiable characteristics. To investigate this concern, data from the 2007 ACS were used to calculate the ratio between the number of children with various combinations of observable demographic characteristics in the NSCH sample and the number of children with those combinations of characteristics in the general population. When the ratio was large and/or the population size was small, some of the identifiable characteristics in the NSCH data file were changed.

- For 87 children, the number of children living in the household (TOTKIDS4) was increased or decreased by one. This change also resulted in perturbation of the variable indicating the age of the sample child relative to other children living in the household (AGEPOS4).
- For 15 children, the number of adults living in the household (TOTADULT3) was increased by one.
- For 4 children whose mothers were not born in the United States, this variable (K11Q30R) was set to “born in the U.S.”
- For 5 children whose fathers were not born in the United States, this variable (K11Q31R) was set to “born in the U.S.”

- For 6 children who were not born in the United States, this variable (K11Q33R) was set to “born in the U.S.” In addition, for 698 children born in and adopted from a foreign country, this variable was also set to “born in the U.S.”
- For 6 children living with a biological, step, foster, or adoptive mother and with no father of any type present, the family structure variable (FAMSTRUCT) was set to “other.”

Analysts interested in working with data that were suppressed or perturbed to protect confidentiality may apply to access unmodified data files through the NCHS Research Data Centers (RDC). These facilities, designed for the researcher outside of NCHS, are located at NCHS headquarters in Hyattsville, Maryland, and at CDC headquarters in Atlanta, Georgia. Unmodified data files can also be made available through Census Bureau RDCs. Data files housed in an RDC may also be accessed remotely via e-mail. Analysts should visit their website at <http://www.cdc.gov/nchs/r&d/rdc.htm> for more information.

Derived Variables

A number of derived variables appear on the data file. The definitions of these variables are provided below.

AGEPOS4—This variable represents the age of the sampled child, relative to the ages of the other children 17 years of age or younger living in the household. Because it is not known if the sampled child was related to the other children living in the household, or if the child has siblings who do not live in the household, or if the child has siblings older than 17 years of age, this variable should not be interpreted as birth order.

AGEYR_CHILD—The child’s age in years was recorded when the child was first identified as the sampled child (which may have been prior to the date when the actual interview was completed). Valid values for age are 0 through 17, where “0” means younger than one year.

CSHCN—This variable is based on the CSHCN Screener (K2Q10, K2Q11, K2Q12, K2Q13, K2Q14, K2Q15, K2Q16, K2Q17, K2Q18, K2Q19, K2Q20, K2Q21, K2Q22, and K2Q23) and indicates whether or not the child has special health care needs.

EDUC_MOMR, EDUC_DADR, and EDUC_RESR—These variables reflecting the highest level of education completed by the mother, father, and respondent, respectively, were derived from data collected in variables K11Q20 through K11Q22.

HISPANIC—This indicator of whether the sampled child is of Hispanic or Latino origin was derived using data collected in variables K11Q01 and K11Q02_OS. Respondents who did not identify a Hispanic ethnicity during administration of K11Q01 but did provide an answer indicating Hispanic ethnicity as part of the verbatim response to the race question (K11Q02_OS) were coded with a value of “1” for the variable HISPANIC.

MARCOH_PAR—This variable indicates the marital/cohabitation status of the child’s parent or parents who live in the household and is based on variables K9Q17A, K9Q17B, K9Q17B_3, K9Q17C,

K9Q17_3 and K9Q17D. Parents living with the child were coded as either married, cohabiting, or neither married nor cohabiting. An additional category reflects that no parents lived in the household.

MARCOH_RESP—This variable indicates the marital/cohabitation status of the NSCH respondent and is based on variables K9Q17A, K9Q17B, K9Q17B_3, K9Q17C, K9Q17_3 and K9Q17D. This variable can be used to impute a value for MARCOH_PAR for cases for which no parents live in the household. This variable is missing for 2,688 cases because one of the following circumstances was true: the respondent was not the mother, and the mother lived in the household with no father present; the respondent was not the father, and the father lived in the household with no mother present; or the respondent was neither the mother nor the father, and both mother and father lived in the household. For those cases, the respondent was asked about the marital/cohabitation status of the child’s parents, not about their own marital/cohabitation status.

MARSTAT_PAR—This variable indicates the legal marital status of the child’s parent or parents who live in the household and is based on variables K9Q17A, K9Q17B, K9Q17C, and K9Q17D. Parents living with the child were coded as either married, separated, divorced, widowed, or never married. An additional category reflects that no parents lived in the household. There is some error associated with this variable: Divorcees who continued to live together after divorce could not be identified as divorced if the household contained a mother and a father, and “never married” includes some cases (4,400 out of 10,392 cases identified as “never married”) for which it is unknown whether the parents were in fact “never” married or simply “not married.”

MARSTAT_RESP—This variable indicates the legal marital status of the NSCH respondent and is based on variables K9Q17A, K9Q17B, K9Q17C, and K9Q17D. This variable can be used to impute a value for MARSTAT_PAR for cases for which no parents lived in the household. This variable is missing for 2,688 cases, for the same reasons as described above for MARCOH_RESP. For those cases, the respondent was asked about the marital status of the child’s parents, not about their own marital status. In addition, as with MARSTAT_PAR, there is some additional error associated with this variable because some divorcees continued to live together after divorce, and because “never married” includes some cases (4,263 out of 9,580 cases identified as “never married”) for which it is unknown whether the respondents were in fact “never” married or simply “not married.”

MSASTAT—This indicator identifying whether or not the household is inside or outside of an MSA was suppressed to protect confidentiality in 16 states.

OTH_LANG—This variable is based on LANG1 and indicates whether the interview was conducted in a language other than English.

PLANGUAGE—This variable was derived from K1Q03 and indicates whether the primary language spoken in the household was not English.

POVERTY_LEVELR—This indicator was created using total household members (K9Q00) and the household income value. If data for either of these two components were missing, refused, or had a “don’t know” response, this measure was assigned a missing value code. The household income value was the actual dollar amount reported by respondents who reported an exact household income

(K11Q51). However, when respondents did not supply a specific dollar amount for household income, it was necessary to go through a series of questions asking respondents whether the household income was below, exactly at, or above threshold amounts (K11Q52 through K11Q59A). If respondents did not complete the income cascade, either because they refused or did not know the answer to one of the cascade questions, this measure was assigned a missing value code. Once an income-to-household-size measure was computed, it was compared with DHHS Federal Poverty Guidelines. More detail about the development of this poverty indicator is available in [Appendix VII](#). Missing values for this poverty indicator were multiply imputed. Details about the development of the imputed values are included in [Appendix XII](#).

RACER, RACEAIAN, RACEASIA, and RACE_HI—These race classification variables were derived from data collected in variables K11Q02X01 through K11Q02X08.

RELATION—Information collected in question K1Q02 regarding the relationship of the respondent to the sampled child has been collapsed into three categories.

SEX—This indicator was created from K1Q01.

SUMMER—This variable indicates that the interview was completed during June, July, or August.

TOTADULT3—The total number of adults in the household was derived by subtracting the total number of children in the household (S_UNDR18) from the total number of persons in the household (K9Q00). As noted previously, this variable was topcoded at 3 or more adults to protect confidentiality.

TOTKIDS4—This variable represents the total number of children 17 years of age or younger living in the household. As noted previously, this variable was topcoded at 4 or more children to protect confidentiality.

Dummy Variables

When respondents were permitted to provide multiple answers for the same question, a variable was created for each possible answer. The values for these new dummy variables are “yes, this answer was given” and “no, this answer was not given.” When respondents could not or did not provide an answer to the question, a value of “don’t know” or “refused” was reported for each of the dummy variables.

- K2Q45D is represented by K2Q45DX01 to K2Q45DX03.
- K4Q28 is represented by K4Q28X01 to K4Q28X04.
- K6Q20B is represented by K6Q20BX01 to K6Q20BX03.
- K6Q31 is represented by K6Q31X01 to K6Q31X03.

Additional Data Notes

The sampled child’s age (AGEYR_CHILD) and the number of children in the household (TOTKIDS4) were updated for a small number of cases during data editing. Data from questions that were skipped due to the previously erroneous age classification were set to missing in error as needed.

On July 5, 2007, several questions related to special health care needs were added for children who were not identified as having special health care needs. These included K2Q12A, K2Q15A, K2Q18A, K2Q21A, K2QTEST1, and K2QTEST2. These questions were intended for methodological research only, and they are not included on the publicly available dataset. Interested researchers may contact NCHS (slaits@cdc.gov) to obtain access to these data.

The question about whether or not doctors provided specific information to address parents' concerns (K6Q11) was intended to be asked for age-eligible children if any concerns were reported at K6Q01 through K6Q09. Due to a CATI programming error implemented on July 5, 2007, the question was asked only when parents reported concerns about the child's learning, development, or behavior (K6Q01) and also reported a concern at K6Q02 through K6Q09. As a result, 5,015 children have missing value for K6Q11.

The age when the child was first fed formula (K6Q42) and the age when the child was first fed anything other than breast milk or formula (K6Q43) could be reported in days, weeks, months, or years. Ages reported in weeks, months, or years were recoded into days (K6Q42R and K6Q43R, respectively).

As the result of a system error, a small number of cases have the variables for time spent reading (K7Q50; 32 cases), time spent computing (K7Q51; 54 cases), or frequency of religious service attendance (K8Q12R; 48 cases) set to "missing in error."

Several questions from the National Survey of Adoptive Parents were included on the NSCH for a subsample of NSCH households. Data from these questions, about school performance (K7Q12 and K7Q13) and about the parent/child relationship (K8Q22 through K8Q26), are not included on the publicly available dataset due to the small sample size (n = 2,022). Interested researchers may contact NCHS (slaits@cdc.gov) to obtain access to these data.

Some respondents initially reported that the mother or father of the selected child does not live in the household (at K9Q11 and K9Q12), but later indicated (at K9Q17B_2 or K9Q17C_2) that the biological mother or father does live in the household. For these cases, the later information was used when deriving the family structure (FAMSTRUCT) and marital/cohabitation variables (e.g., MARCOH_PAR). This initial reporting error occurred for mothers in 189 cases and for fathers in 533 cases. The logic for asking parent-specific questions later in the survey (such as health status of each parent and education level of each parent) was based on the initially erroneous K9Q11 and K9Q12 responses. Answers to these parent-specific questions were set to "missing in error" as appropriate.

The text fill at K9Q18 was revised mid-quarter in Q2/2007 to ask about the respondent's spouse or partner rather than the child's mother or father because this question could be misinterpreted by respondents in step-families. Data collected prior to this revision were not suppressed during editing because it is unknown which respondents misunderstood the question.

There are 67 NSCH partial completes where the NIS interview was completed. Where applicable, responses from NIS demographic questions have been filled into the corresponding NSCH questions.

Procedures for Developing Sampling Weights

This section provides a nontechnical overview of the weighting procedures for the NSCH sample. A more detailed and technical description can be found in [Appendix III](#).

For the NSCH, a single weight was generated for all analyses. The steps to calculate this weight consist of the calculation of a base sampling weight, the development of full sample weights from quarterly sample weights, an adjustment for nonresolution of telephone numbers, adjustments for nonresponse, an adjustment for subsampling of age-eligible children, an adjustment for multiple telephone lines, an adjustment for noncoverage of children in households without landline telephones, and a raking adjustment to external control totals.

Base Weight

The goal of the NSCH was to complete approximately 1,750 interviews in each state over five calendar quarters of data collection. The total number of telephone lines needed to obtain this number of completed interviews was estimated. Enough NIS sample was subsequently selected for most estimation areas to obtain the required number of completed cases for the NSCH for each quarter of data collection. Some estimation-area samples contained too few telephone numbers in the NIS sample to obtain the desired number of completed cases. In these areas, additional telephone numbers were randomly selected to reach the NSCH targets.

The telephone lines selected for screening for the NSCH represent a random sample in each geographic area of all possible telephone lines in non-cellular banks of telephone numbers containing at least one residential-listed number. The probability that any given telephone line will be selected from this population of possible telephone lines can be calculated by dividing the number of telephone lines selected for the study by the total number of possible telephone lines in a given area.

Each telephone line selected for the NSCH represented some larger number of telephone lines in that geographic area. This number can be calculated as the inverse of the probability of selection for any telephone line. This number is the base weight that is associated with each completed household interview in that geographic area.

In computing the base weight, a special adjustment was applied to accommodate the subsampling of the NSCH sample that occurred for Q1/2007. The final released sample for the NSCH in Q1/2007 was a subsample of the cases initially flagged for release. Therefore, the base weights for this banked sample were equal to the normal base weight times a factor equal to the inverse of the probability of being included in the subsample.

Computing Full Sample Weights from Quarterly Weights

In this step, all five quarterly samples were combined such that the samples from all quarters jointly represent the full population. Because the base weights were calculated for each quarter separately, the sum of the base weights in each quarter represents the full population for a geographic area. For each area, the full-sample weights were computed from quarterly base weights by applying composition factors proportional to the number of sampled and released telephone numbers in a quarter, adjusted for the design effect due to any differential sampling rates within the quarter.

First Form of Nonresponse: Unknown Household Status

When the selected telephone lines were called, three results were possible:

- It was determined that the telephone line belonged to a household.
- It was determined that the telephone line was not a working residential number but was a business number or a nonworking number.
- The status was undetermined because the telephone rang without being answered, the person answering the telephone hung up immediately, or the telephone-answering device did not indicate whether the telephone line belonged to a household.

This last category includes some household telephone lines, but the exact number of household telephone lines in this category is unknown. The households with resolved telephone numbers must represent the households in this unknown category. The size of this nonresponse adjustment is based on the observed size of the first two categories. The proportion of households in the unknown category is assumed to be the same as the proportion of households among all resolved telephone numbers. This adjustment varies based on socioeconomic and demographic characteristics of the population under a telephone exchange and whether the telephone line was directory listed. Based on the frequency of the nonresponse in a given adjustment cell, compensation is made for this nonresponse by proportionately increasing the weights for those households with resolved telephone numbers, so that the households with resolved telephone numbers represent the households in the unknown category.

Second Form of Nonresponse: Unknown Household Eligibility

When a household has been identified, three results are possible:

- It is determined that the household includes an age-eligible child.
- It is determined that the household does not include a child and is not eligible.
- The age-eligibility screener is not completed, and the eligibility of the household is unknown.

This last category includes some age-eligible households. However, the exact number of age-eligible households in this category is unknown. The households known to be age-eligible must represent the age-eligible households in this unknown category. This proportional adjustment is the second unit nonresponse adjustment. The size of the adjustment is based on the observed size of the first two categories. The proportion of age-eligible households in the unknown category is assumed to be the same as the proportion of age-eligible households among all households where the screening interview for the presence of children was completed. Similar to the previous step, this adjustment varies based on socioeconomic and demographic characteristics of the population under a telephone exchange and whether the telephone line was directory listed. Based on the frequency of nonresponse to the age-eligible screening interview in an adjustment cell, compensation is made for this nonresponse by proportionately increasing the weights for those households known to be age-eligible in that cell, thus representing the age-eligible households in the unknown category. The nonresponse adjustment for age screening was made within each state.

Adjustment for Households with More than One Child

The unit of analysis now rotates to the child level from the household level. For households with more than one age-eligible child, only one child was selected randomly per household to be the subject of the NSCH interview. The randomly selected child represents all of the children in the household. Therefore, the sampling weight for this randomly selected child must be increased to reflect the fact that this child represents multiple children in that household. This adjustment multiplies the nonresponse adjusted household weight by the number of children in the household.

Third Form of Nonresponse: Sampled Children for Whom an Interview Is Not Completed

When a child has been randomly selected (i.e., sampled), two results are possible:

- An interview is completed or sufficiently partially completed.
- An interview is not completed.

The completed child interviews must also represent the children who were sampled but for whom an interview was not completed. This proportional adjustment is the third unit nonresponse adjustment. The size of the adjustment is based on the size of the two categories and is calculated simply as the ratio of the weighted total number of sampled children to the weighted number of completed interviews. In other words, based on the frequency of nonresponse among sampled children with certain socio-demographic characteristics in a given state, compensation is made for this nonresponse by proportionately increasing the weights for those interviews that could be completed in the adjustment cell. The completed interviews, therefore, also represent the sampled children with incomplete interviews.

Adjustment for Multiple Telephone Landlines

Among the households that complete the interview, some will report more than one landline telephone for home use (excluding lines used only for fax or computer). If a household has multiple telephone lines for home use, this household has a greater chance of being included in the survey than does a household with only a single telephone line. An adjustment to the weight is necessary to compensate for their multiple chances of selection. The adjusted child-level weight from the previous step was further adjusted by dividing the weight by the number of telephone landlines for home use.

Adjustment for Noncoverage of Children in Non-Landline Telephone Households

The NSCH did not select households without a landline telephone at the time of the survey. To compensate for the noncoverage of children in non-landline telephone households, an adjustment based on whether the child's household experienced an interruption in telephone service was applied. Households with interrupted telephone service were targeted in the weighting process because there is evidence that households with telephones at the time of the survey, but with interruptions in telephone service during the year, are more similar to households with no telephone service than are households with uninterrupted telephone service during the year (28-30). Therefore, noncoverage of households without a telephone can be somewhat compensated for by proportionately increasing the weights for those interviews that could be completed in households with interrupted service. In this way, completed

interviews in households with interrupted service also represent the incomplete interviews in households without telephone service at the time of the interview.

To apply this adjustment, two groups were defined within each state depending on whether the household had an interruption in landline telephone service of more than one week or not during the past 12 months. The size of the adjustment was based on state-level estimates of the proportion of children in households with a landline telephone, derived based upon 3 years of National Health Interview Survey data.

Raking Adjustment

Despite the weighting efforts and the nonresponse adjustments, the estimated number of children is unlikely to perfectly match known population totals. Any discrepancies are likely to be due to random sampling error and nonrandom response biases. Raking adjusts the weights to match population control totals for key socio-demographic information obtained from external sources.

For the NSCH, the independent source was the 2006 and 2007 American Community Survey (ACS). The total counts of children by state were obtained from the 2007 ACS, which corresponds to July 1, 2007. The mid-point of the NSCH data collection periods varied from state to state but, for deriving control totals, July 1, 2007 was used as the reference time point for all states. The counts by different subgroups within each state were derived by first estimating proportions in subgroups and then applying the estimated proportions to the total counts of children in the state. The ACS 2006 and 2007 files were combined to produce stable estimates of proportions by state for all but one variable. For the number of children in each state living in metropolitan statistical areas (MSAs), the proportional distribution was obtained by combining three years (2005-07) of Current Population Survey (CPS) data. ACS and CPS estimates were restricted to the noninstitutionalized population of children, which includes regular households and non-institutional group quarters.

The raking adjustment was done in two steps: a short initial step and then the primary raking step with all essential control totals. At the initial step, the raking adjustment was done using various categories of the following marginal totals within each state and DC:

- Number of male and female children within each of four age groups
- Number of children in MSAs and non-MSA areas

The purpose of this initial step is to ensure that the input weights to the primary raking step were adjusted for MSA status. To accommodate other more important raking dimensions, the MSA dimension was not included in the primary raking step.

At the primary raking step, the adjustment was done using various categories of the following marginal totals within each state and DC:

- Number of male and female children within each of three age groups
- Number of children in various nonoverlapping race and ethnicity categories, where the number of categories varied by state

- Number of children in households that have a highest reported level of education within each of three nonoverlapping categories
- Number of children in households that have a household income in each of five nonoverlapping categories
- Number of children in households with one child, in households with two children, and in households with three or more children.

For raking purposes only, all missing variables used in the weighting process were imputed using a hot deck imputation procedure. The donors were selected by forming imputation classes for each imputed variable.

The various adjustments occasionally made some weights substantially larger than other weights. These extremely large weights were truncated to prevent a small number of cases with large weights from having undue influence on estimates. After any extreme weights were trimmed, the weights were re-raked. This process was iterated until there were no extreme weights after raking. The raking and trimming process yields a final weight for each child with a complete NSCH interview.

Quality Control

NORC staff compared the formulas for the weights and adjustments developed by the sampling statistician with the actual weights and adjustments constructed by the statistical programmer. An independent check was performed on the programmer's implementation of the statistician's weighting specifications.

Estimation and Hypothesis Testing

The NSCH data were obtained through a complex sample design involving unequal probabilities of selection of children within households and stratification of households within states. To produce estimates that are representative of children nationally and within each state, sampling weights must be used. As described earlier, a single sampling weight (NSCHWT) has been developed for the NSCH. This weight should be used for both national and state-level analyses.

Interpretation of Weighted Estimates

Estimates based on the sampling weights generalize only to the population of US noninstitutionalized children ages 0 through 17 years at the time of the interview. These estimates do not generalize to the population of parents, the population of mothers, or the population of children's health care providers.

Two examples may help make this distinction clearer. Weighted estimates based on K8Q11 can be interpreted as the proportion of children whose families regularly eat meals together, but should not be interpreted as the proportion of families who regularly eat meals together. Similarly, weighted estimates based on K8Q30 can be interpreted as the proportion of children whose parents are coping well with the demands of parenthood, but should not be interpreted as the proportion of parents who are coping well.

Variables Used for Variance Estimation

Because of the complex design of the NSCH, the interviewed cases have unequal weights. Therefore, statistical software programs that assume simple random sampling will most often compute standard errors that are too low. Tests of statistical hypotheses may then suggest statistically significant differences or associations that are misleading. However, computer programs are available that provide the capability of variance estimation for complex sample designs (e.g., SUDAAN, Stata, WesVar). To provide the user with the capability of estimating the complex sample variances for the NSCH data, we have provided stratum identifiers and primary sampling unit (PSU) codes on the data files. These variables and the sample weights are necessary to properly calculate variances.

The strata identified on the publicly released data set are not identical to the strata used to draw the sample. In states with multiple estimation areas, independent samples were selected from each estimation area in proportion to the total number of households with children in each estimation area. Therefore, these estimation areas should be considered strata for variance estimation. However, disclosure of the specific estimation area for each child (even if the code were scrambled) could increase the risk of disclosure of a child's or respondent's identity. For example, the estimation area with the lowest frequency of responses in New Jersey would be readily identifiable as Newark. In the absence of estimation area-specific identifiers, data users should use the state identifier (STATE) as the stratum identifier. By using the state identifier rather than the suppressed estimation area identifier, the standard errors for national and state estimates with key variables are affected only slightly, and not in a consistent direction. The PSU for the NSCH is the household, represented on the data sets by the unique household identifier, IDNUMR.

The overall number of persons in this survey is sufficient for most statistical inference purposes. However, analyses of some rare responses and analyses of subclasses can lead to estimators that are unreliable. Small sample sizes used in the variance calculations may also produce unstable estimates of the variances. Consequently, these analyses require that the user pay particular attention to the variability of estimates of means, proportions, and totals.

Variance Estimation Using SUDAAN or STATA

Standard errors of NSCH estimates can be obtained using the Taylor series approximation method, available in software such as SUDAAN, SAS, and STATA. The state is the stratum variable and the household is the PSU variable.

The simplifying assumption that PSUs have been sampled with replacement allows most complex survey sample design computer programs to calculate Taylor series standard errors in a straightforward way. This method requires no recoding of design variables, but is statistically less efficient (and therefore more conservative) than some other methods because the PSU unit is treated as being sampled with replacement within the stratum unit. For SUDAAN, the data file needs to be sorted by stratum (STATE) and PSU (IDNUMR) prior to invoking any procedures. The following example lists SUDAAN design statements that can be used to analyze the data:

- PROC ... DESIGN = WR;
- NEST STATE IDNUMR;

- WEIGHT NSCHWT;

For STATA, the following design statements are used:

- svyset strata STATE
- svyset psu IDNUMR
- svyset pweight NSCHWT
- svyset

Other variance estimation procedures are also applicable to the NSCH. Specifically, the jackknife method with replicate weights and the bootstrap resampling method with replicate weights can also be used (via software such as WesVar) to obtain standard errors that fully reflect the impact of the weighting adjustments on standard errors.

Variance Estimation for Subsets of the Data

Most analyses of the NSCH data will focus on specific population subgroups, such as children in only one state or children who live in poverty. Some analysts will therefore be tempted to delete all records outside of the domain of interest in order to work with smaller data files and run computer jobs more quickly. This procedure of keeping only select records and deleting other records is called subsetting the data. Subsetted data that are appropriately weighted can be used to generate correct point estimates (e.g., estimates of population subgroup frequencies or means), but most software packages that analyze complex survey data will incorrectly compute standard errors for subsetted data. When complex survey data are subsetted, the sample design structure is often compromised because the complete design information is not available. Subsetting the data can delete important design information needed for variance estimation.

The NSCH was designed to provide independent data sets for each of the 50 states and DC. Subsetting the survey data to a particular state does not compromise the design structure of the survey. That is, standard errors calculated in SUDAAN for a particular state will not be affected if the data set has been subsetted to that particular state.

However, subsetting to specific population subgroups (within or across states) can result in incorrect standard errors. For example, subsetting the data to children who live in poverty within a specific state will result in incorrectly calculated standard errors. Typically, the standard errors for subsetted data will be inflated, resulting in a higher probability of type II error (i.e., failing to detect significant differences that do in fact exist). SUDAAN has a SUBPOP option that allows the user to target specific subpopulations for analysis while retaining the full unsubsetted data set that includes the full sample design information. Analysts interested in specific population subgroups must use SUBPOP rather than subsetting the data sets.

Weighted Frequencies, Prevalence Estimates, and Standard Errors

Weighted state-specific frequencies of the number of children with excellent or very good health (as assessed by the respondent) appear in [Appendix XIII](#). Prevalence estimates and standard errors are

also provided. Analysts may wish to replicate this table to determine if they are using the weights correctly.

Weighted frequencies, prevalence estimates, and standard errors for other survey measures are available from the Data Resource Center for Child and Adolescent Health. This online center is led by the Child and Adolescent Health Measurement Initiative (CAHMI) at the Oregon Health and Science University, and is supported through a cooperative agreement with MCHB. The data resource center is accessible at <http://www.nschdata.org> or <http://www.childhealthdata.org>.

Guidelines for Data Use

With the goal of mutual benefit, NCHS requests that recipients of data files cooperate in certain actions related to their use.

Any published material derived from the data should acknowledge NCHS as the original source. The suggested citation, “Data Source: National Center for Health Statistics and Maternal and Child Health Bureau, National Survey of Children’s Health, 2007” should appear at the bottom of all tables. Published material derived from the data should also include a disclaimer that credits any analyses, interpretations, or conclusions reached to the author and not to NCHS, which is responsible only for the initial data. Consumers who wish to publish a technical description of the data should make a reasonable effort to ensure that the description is not inconsistent with that published by NCHS.

CIPSEA and the Public Health Service Act (Section 308d) provide that these data collected by NCHS may be used only for the purpose of health statistical reporting and analysis. Any effort to determine the identity of any reported case is prohibited by these laws. NCHS takes extraordinary measures to assure that the identity of survey subjects cannot be disclosed. All direct identifiers, as well as any characteristics that might lead to identification, have been omitted from the data set. Any intentional identification or disclosure of a person or establishment violates the assurances of confidentiality given to the providers of the information. Therefore, users must:

- Use the data in this data set for statistical reporting and analysis only;
- Make no use of the identity of any person discovered, inadvertently or otherwise, and advise the Director, NCHS, of any such discovery (301-458-4500);
- Not link this data set with individually identifiable data from any other NCHS or non-NCHS data sets.

Use of the data set signifies users’ agreement to comply with the above-stated statutory-based requirements.

Further Information

Data users can obtain the latest information about SLAITS by periodically checking the SLAITS website at <http://www.cdc.gov/nchs/slaits.htm>. This site features downloadable data files and documentation for SLAITS modules, as well as important information about any modifications and

updates to data and/or documentation. Data users will also find current contact information if you have any additional questions. Data users with questions may also send e-mail to slaits@cdc.gov.

Researchers may also wish to join the SLAITS electronic mail listserv. To subscribe or unsubscribe, visit <http://www.cdc.gov/nchs/about/major/slaits/slaitslistserv.htm> and follow the directions listed. The listserv has approximately 1,000 subscribers around the world who use SLAITS data or are interested in SLAITS. Subscribers periodically receive e-mail containing news about SLAITS surveys (e.g., new releases or modifications to existing data), publications, or related conferences. The listserv is moderated and listserv membership is private.

For more information on CDC, you may contact CDC's Information Contact Center (CDC-INFO) in English or Spanish by calling (800) CDC-INFO [800-232-4636] or e-mailing cdcinfo@cdc.gov. Persons with hearing impairment may contact CDC-INFO with a TTY machine at (888) 232-6348. The CDC-INFO fax machine line is (770) 488-4760. Please note, however, that CDC-INFO cannot respond to questions about individual medical cases, provide second opinions, or make specific recommendations regarding therapy. These issues should be addressed directly with personal health care providers.

References

1. Ireys HT, Nelson RP. New federal policy for children with special health care needs: implications for pediatricians. *Pediatrics* 90:321–7. 1992.
2. Health Resources and Services Administration. Maternal and Child Health Bureau Strategic Plan: FY 2003–07. Available at: <http://mchb.hrsa.gov/about/stratplan03-07.htm>.
3. Ezzati-Rice TM, Cynamon M, Blumberg SJ, Madans JH. Use of an existing sampling frame to collect broad-based health and health-related data at the state and local level. Federal Committee on Statistical Methodology Statistical Policy. Working Paper Series 30:283–92. 2000.
4. Blumberg SJ, Olson L, Osborn L, et al. The design and operation of the National Survey of Early Childhood Health, 2000. National Center for Health Statistics. *Vital and Health Statistics* 1(40). 2002.
5. Blumberg SJ, Olson L, Frankel M, Osborn L, Becker CJ, Srinath KP, Giambo P. Design and operation of the National Survey of Children with Special Health Care Needs, 2001. National Center for Health Statistics. *Vital and Health Statistics* 1(41). 2003.
6. Blumberg SJ, Olson L, Frankel MR, Osborn L, Srinath KP, Giambo P. Design and operation of the National Survey of Children's Health, 2003. National Center for Health Statistics. *Vital and Health Statistics* 1(43). 2005.
7. O'Connor KS, Osborn L, Olson L, Blumberg SJ, Frankel MR, Srinath KP, Giambo P. Design and operation of the National Asthma Survey. National Center for Health Statistics. *Vital and Health Statistics* 1(46). 2008.
8. Blumberg SJ, Welch EM, Chowdhury SR, Upchurch HL, Parker EK, Skalland BJ. Design and operation of the National Survey of Children with Special Health Care Needs, 2005-06. National Center for Health Statistics. *Vital and Health Statistics* 1(45). 2009.
9. van Dyck P, Kogan MD, Heppel D, et al. The National Survey of Children's Health: A new data resource. *Maternal and Child Health Journal* 8:183–8. 2004.
10. NORC at the University of Chicago. 2007 Methodology Report: National Immunization Survey. Chicago, IL: NORC. 2008.
11. Zell ER, Ezzati-Rice RM, Battaglia MP, Wright RA. National immunization survey: The methodology of a vaccination surveillance system. *Public Health Reports* 115:65-77. 2000.
12. Kalsbeek WD, Agans RP. Sampling and weighting in household telephone surveys. In JM Lepkowski, C Tucker, JM Brick, E de Leeuw, L Japec, PJ Lavrakas, MW Link, RL Sangster (eds.). *Advances in Telephone Survey Methodology*. Hoboken, New Jersey: John Wiley & Sons, Inc., pp. 29-55. 2008.

13. Tucker C, Casady RJ, and Lepkowski, JM. A Hierarchy of List-Assisted Stratified Telephone Sample Design Options. *Proceedings of the Survey Research Methods Section*. Alexandria, VA: American Statistical Association, pp. 982–987. 1993.
14. Blumberg SJ, Luke JV, Cynamon ML, Frankel MR. Recent trends in household telephone coverage in the United States. In JM Lepkowski, C Tucker, JM Brick, E de Leeuw, L Japec, PJ Lavrakas, MW Link, RL Sangster (eds.). *Advances in Telephone Survey Methodology*. Hoboken, New Jersey: John Wiley & Sons, Inc., pp. 56-86. 2008.
15. Bethell CD, Read D, Stein RE, et al. Identifying children with special health care needs: Development and evaluation of a short screening tool. *Ambulatory Pediatrics* 2:38–48. 2002.
16. American Academy of Pediatrics. The medical home [policy statement]. *Pediatrics* 110:184–6. 2002.
17. Glascoe FP. Parents' evaluation of developmental status: a method for detecting and addressing developmental and behavioral problems in children. Nashville, TN: Ellsworth & Vandermeer Press, LLC. 1997.
18. Blumberg SJ, Carle AC, O'Connor KS, Moore KA, Lippman LH. Social competence: development of an indicator for children and adolescents. *Child Indicators Research* 1:176-197. 2008
19. Abidin RR. Parenting Stress Index: A measure of the parent-child system. In CP Zalaquett, RJ Wood (eds.), *Evaluating stress: A book of resources* (pp. 277–91). Lanham, MD: Scarecrow Press, Inc. 1997.
20. Easterbrooks MA, Goldberg WA. Toddler development in the family: Impact of father involvement and parenting characteristics. *Child Development* 55:740–52. 1984.
21. Fields JM, Smith KE. Poverty, family structure, and child well-being: Indicators from the SIPP. Population Division Working Paper No. 23. U.S. Bureau of the Census, Population Division. 1998. Available at: <http://www.census.gov/population/www/documentation/twps0023/twps0023.html>.
22. Sampson RJ, Raudenbush SW, Earls F. Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science* 277:918–24. 1997.
23. De leeuw ED, Callegaro M, Hox J, Korendijk E, Lensvelt-Mulders G. The influence of advance letters on response in telephone surveys. *Public Opinion Quarterly* 71:413-443. 2007.
24. The American Association for Public Opinion Research. Standard definitions: Final dispositions of case codes and outcome rates for surveys. 5th edition. Lenexa, KS: AAPOR.. 2008.
25. Kuczmarski RJ, Ogden CL, Grummer-Strawn LM, et al. CDC growth charts: United States. Advance data from vital and health statistics; no. 314. National Center for Health Statistics. 2000.

26. CDC. A SAS program for the CDC growth charts. Available at:
<http://www.cdc.gov/nccdphp/dnpa/growthcharts/resources/sas.htm>.
27. Akinbami LJ, Ogden CL. Childhood overweight prevalence in the United States: The impact of parent-reported height and weight. *Obesity*. Forthcoming.
28. Brick JM, Waksberg J, Keeter S. Using data on interruptions in telephone service as coverage adjustments. *Survey Methodology* 22:185–197. 1996.
29. Davern M, Lepkowski J, Call KT, et al. Telephone service interruption weighting adjustments for state health insurance surveys. *Inquiry* 41:280–290. 2004.
30. Frankel MR, Srinath KP, Hoaglin DC, et al. Adjustments for non-telephone bias in random-digit-dialing surveys. *Statistics in Medicine* 22:1611–1626. 2003.
31. Landreth A. SIPP advance letter research: cognitive interview results, implications, & letter recommendations. U.S. Census Bureau, Center for Survey Methods Research. 2001. Available at:
<http://www.census.gov/srd/papers/pdf/sm2001-01.pdf>.
32. Rubin DB. Multiple imputation for nonresponse in surveys. New York: John Wiley. 1987.
33. Raghunathan TE, Lepkowski JM, Van Hoewyk J, Solenberger P. A multivariate technique for multiply imputing missing values using a sequence of regression models. *Survey Methodology* 27:75–85. 2001.
34. Schafer JL. (1997). Analysis of incomplete multivariate data. London: Chapman & Hall. 1997.
35. Pedlow S, Luke JV, Blumberg SJ. Multiple imputation of missing household poverty level values from the National Survey of Children with Special Health Care Needs, 2001, and the National Survey of Children's Health, 2003. Hyattsville, MD: National Center for Health Statistics. 2007. Available at: http://www.cdc.gov/nchs/data/slats/mimp01_03.pdf.

Appendix I

Sample Design

The basic design objective of the NSCH sample was to interview a sample of 1,700 children younger than 18 years of age in each state and the District of Columbia. The sample was selected by identifying households with children under the age of 18. If only one child lived in the household, that child was the target of the interview. If more than one child was present, one child was randomly selected as the target.

National Immunization Survey sample

With the exception of the NSCH-only augmentation sample, the sample of households selected to be screened for the NSCH was predominately a subsample of the households screened for the National Immunization Survey (NIS), a continuous list-assisted random-digit-dialed (RDD) survey. Starting in 2007, the base NIS estimation areas included 56 regions (50 state or “rest of state” areas plus 6 grantee urban areas). The six grantee urban areas are: Chicago, IL; Philadelphia, PA; New York City, NY; Bexar County, TX; Houston County, TX; and Washington, DC. Also starting in 2007, state immunization programs could identify cities or counties of interest to be oversampled. Eight of those regions, which may change annually, were selected and added to the base NIS estimation areas to equal a total of 64 estimation areas. For more detail on the NIS sample design, readers are encouraged to obtain the *2007 Methodology Report* for the NIS (10).

Associating telephone numbers with estimation areas

To draw a sample of telephone numbers in an NIS estimation area, a list of all telephone numbers that belong to that area must be compiled. For some estimation areas, this step is straightforward. For example, when the estimation area is a state, the list would consist of all telephone numbers within the central-office codes that are in service in the area codes assigned to that state. (Combined, an area code and a central-office code form a “prefix area.” For example, 312-555-xxxx is the prefix area corresponding to the 555 central office in the 312 area code.)

For other estimation areas, however, this step encounters a number of complications. When the estimation area is a city or county, some prefix areas may cover part of the estimation area and part of an adjacent estimation area. In such situations, the NIS applies a majority rule: If at least 50% of the directory-listed households in a prefix area fall inside an estimation area, the prefix area is assigned to that estimation area.

Drawing the initial National Immunization Survey sample

The sampling frame for an estimation area consists of banks of 100 consecutive telephone numbers within the prefix areas assigned to the estimation area. Banks that contain only wireless (i.e., cell phone) telephone numbers are excluded from the frame. Banks that contain zero directory-listed residential telephone numbers are also excluded from the frame because they have very little chance of containing working residential numbers.

To exclude banks that contain zero directory-listed residential telephone numbers, the GENESYS Sampling System (a proprietary product of Marketing Systems Group (MSG)) uses a file of directory-listed residential numbers from InfoUSA. The result is a file that lists the remaining banks (the “1+ working banks”). From the 1+ working banks, a random sample of complete ten-digit telephone numbers is drawn for each quarter in such a way that each number has a known probability of selection that is equal for all numbers in the same estimation area. Within each estimation area, the NIS sample is then segmented into replicates, or representative subsamples, with each replicate containing sample telephone numbers from each of the estimation areas. Segmenting the sample into replicates allows the release of telephone numbers over time in a controlled manner.

Updating the National Immunization Survey sampling frame

The set of telephone banks with at least one directory-listed residential telephone number changes over time. As a result, the NIS sampling frame needs to be updated on a quarterly basis. Area-code splits produce additional changes to the sampling frame. MSG maintains a separate sampling frame for each estimation area. Each quarter, MSG examines the database to determine whether any currently included banks should be assigned to different estimation areas and to assign newly included banks to estimation areas. The rules for assignment are the same as in the initial definitions of the estimation areas.

Once all modifications have been made to the database, a number of checks ensure that all changes have been applied correctly and that the new database produces samples consistent with those produced prior to the changes. These checks compare the numbers of active banks and RDD-selectable lines in each estimation area before and after the update. In parallel, the numbers of exchanges assigned to each estimation area before and after the update are compared. Small changes are expected, because new banks are put into service as new numbers are assigned. In the event of a major discrepancy in any of these checks, MSG is notified of the difference and asked to provide documentation of the reasons for the change.

Preparation of the Sample

Coordinated management of the sample follows a sequence of steps. The initial quarterly sample for each estimation area is divided into replicates. Before a replicate is loaded into the CATI system, several stages of processing remove as many businesses and nonworking numbers as possible. Telephone numbers on the NIS’s “Do Not Call List” are not loaded into the CATI system. Also, at each quarter, any duplicate telephone numbers (i.e., numbers that have appeared in the sample in any of the three prior quarters) are identified and excluded from dialing. For the 2007 NSCH, 52.7% of the telephone numbers in released replicates were loaded into the CATI system for calling. The remaining 47.3% of the telephone numbers in the released sample were resolved prior to dialing.

Forming National Immunization Survey sample replicates

The NIS divides the sample in each estimation area into 26 representative subsamples or replicates. This procedure permits smoother release of the sample (at the rate of one or two replicates per week) for each estimation area separately, as needed. Toward the end of the quarter, half-size replicates allow tighter control over the total amount of sample released. The aim is to produce an even distribution of work in the telephone center over the course of each quarter of data collection.

Removing business and nonworking numbers

Over two-thirds of all selected telephone numbers are typically businesses or unassigned. It would be incredibly inefficient to require the interviewers to dial and classify all of these numbers. To prevent that potential expense, the NIS uses another MSG product (a companion to the GENESYS Sampling System) to quickly and accurately reduce the size of this task.

First, the selected sample is matched against a GENESYS file containing telephone numbers that are directory-listed in a business Yellow Pages and are not directory-listed in a residential White Pages. Any business numbers so identified are removed from the sample. Second, numbers listed in residential White Pages are identified and temporarily set aside.

Third, a hardware system (GENESYS-IDplus) screens the remaining sample to remove a portion of the nonworking numbers. Using personal computers with special hardware and software, this system (the “autodialer”) automatically dials the telephone numbers to detect non-working numbers, which are indicated by the familiar tri-tone signal for out-of-service numbers, by an extended period of silence, or by continuous noise on the line. Sometimes non-working numbers ring one or more times before the tri-tone occurs; GENESYS-IDplus permits numbers to ring two times before hanging up. On a national basis, 15% to 20% of the dialed numbers ring and are answered. To minimize the number of answered calls, the system is used only during the day, when household members are less likely to be at home. Calls that are answered are routed to an attendant, who says, “Sorry, I must have dialed the wrong number.”

Finally, the directory-listed residential numbers are combined with the numbers that were not removed by the autodialer to produce the NIS sample for the telephone center. The numbers removed within released replicates are themselves considered released. They are also considered pre-screened and assigned disposition codes indicating that they are resolved, non-residential numbers.

Ported wireless telephone numbers

A significant recent development in the telecommunications industry is the FCC regulation on portability. Local number portability allows wireless telephone customers to switch from one company to another while retaining the same telephone number. There are three ways in which consumers can take advantage of the new number portability provisions: 1) wireless-to-wireless, 2) wireless-to-landline, and 3) landline-to-wireless.

Wireless telephone numbers are not knowingly included in the RDD sampling frame. Therefore, the first way to make a number portable does not impact the RDD sampling strategy. The second way could result in the exclusion of a small number of households with landline telephone service from the RDD sampling frame. The third way – the porting of landline numbers to wireless service providers – creates the possibility of inadvertently including wireless telephone numbers in the RDD samples.

FCC rules (implementing the Telephone Consumer Protection Act of 1991) bar automated calls to wireless telephone numbers. To pre-identify landlines that have been ported to wireless telephones, the selected NIS sample is matched to the NeuStar database, which contains the national list of ported telephone numbers. Details on the database can be found at <http://www.tcpacompliance.com>. Each quarterly sample is compared to the database and the ported numbers flagged accordingly. The flagged

numbers are assigned an out-of-scope disposition code and are not called. The numbers in released replicates are also matched to the NeuStar database on a daily basis to identify any new ports that have not already been finalized within the telephone center. If a number is dialed in the NIS and found to have reached a wireless telephone number, the call is immediately terminated and classified as out of scope without seeking an interview.

“Do not call” requests

A file is maintained by NORC containing telephone numbers of people who have previously been contacted for the NIS and have requested that they not be called again. Each quarter’s sample is compared with this file, and numbers contained in the “Do Not Call List” are not included in the quarterly sample of numbers loaded into the CATI system.

Duplicate telephone numbers

Because of the repeated quarterly sampling operations in each estimation area, some telephone numbers were selected more than once. To avoid any respondent problems created by recontacts for the same survey, a further sample preparation step identified duplicate numbers. Each quarterly NIS sample file was compared with all sample files for the three previous quarters, and the duplicate numbers were excluded. Thus, the quarterly samples were essentially selected by a method of without-replacement sampling. However, analysts are reminded to invoke with-replacement sampling in SUDAAN or Stata for accurate variance estimation.

Obtaining Addresses for Advance Letters

To obtain addresses that correspond to telephone numbers in the sample, the numbers for each replicate are sent to a vendor (TARGUSinfo) who maintains a large and frequently updated database of over 160 million residential and business telephone numbers, including unpublished telephone numbers. Sources for the data include call centers and companies in the telecommunications, consumer goods, insurance, and credit industries.

After the sample has been prepared, the use of this vendor’s reverse-match system yielded addresses for about 58.8% of the telephone numbers loaded into the CATI system. Advance letters were sent to these addresses. The mailing was approximately 10 days, or two weekends, prior to the time when the telephone numbers in the corresponding replicates were scheduled to be called.

Appendix II

Banked Sample

In the usual SLAITS operation, households either screen as ineligible for the NIS or as NIS-eligible and complete the NIS interview. These households then move directly – on the same call, where feasible – on to the SLAITS portion of the CATI instrument for SLAITS screening and interviewing. However, in Q1/2007 the CATI instrument was not finalized at the time the NIS was being fielded; therefore, households could not move on to NSCH screening and interviewing directly upon finishing the NIS.

As in other quarters, in Q1/2007 a portion of the NIS sample was flagged for the NSCH even though NSCH screening and interviewing could not be done in Q1/2007. To take advantage of the NIS screening that was done in Q1/2007, a subsample of the initially-flagged NSCH sample was drawn, using the outcomes of the NIS dialing operation as stratifiers to increase efficiency.

First, every NSCH-flagged case that finished the NIS (either as NIS-ineligible or as NIS-eligible with a completed NIS interview) was included in the subsample with certainty. These cases were dialed for NSCH screening and interviewing between May and October, 2007.

Next, the NSCH-flagged cases that did not finish the NIS (i.e., they did not complete the NIS age-eligibility screener or did not complete the NIS interview) were rescreened by GENESYS-IDplus in November of 2007 to identify cases that could be screened out as businesses or non-working telephone numbers. Any case that was screened out by GENESYS-IDplus was included in the subsample with certainty. However, it was never dialed for NSCH screening and interviewing.

Finally, 10,000 cases were selected from among the NSCH-flagged cases that did not finish the NIS and were not screened out by GENESYS-IDplus. The cases were selected using state and NIS outcome as strata. The 10,000 cases were allocated to each state in proportion to the number of cases initially flagged for the NSCH in the state. (For example, if 1/40th of Q1/2007 initially-flagged NSCH cases fell within a particular state, then that state was allocated 1/40th of the 10,000 cases, or 250 cases.) Within each state, the cases were grouped into four strata according to the outcome of the NIS dialing effort in Q1/2007:

- Group 1: Screened for NIS age-eligibility
- Group 2: Identified as a known household by the NIS, but not screened for NIS age eligibility
- Group 3: Final NIS disposition of “likely household” or “answering machine”
- Group 4: Final NIS disposition of “non-contact,” “non-residential,” or “non-working telephone number”

Within each state, cases from Group 1 were selected with certainty, and cases in Groups 2, 3, and 4 were selected at an equal rate. This rate varied by state so as to yield the appropriate total number of cases selected in each state based on the state-level allocation of the 10,000 cases. Once the 10,000 cases

were chosen in this way, the cases were dialed for NSCH screening and interviewing from December 2007 to April 2008.

In this way, the final Q1/2007 NSCH sample consisted of a subsample of the initially-flagged Q1/2007 NSCH sample, in three pieces:

- 1) Cases that finished the NIS either as NIS-ineligible or NIS-eligible with a completed NIS interview. These cases were included in the subsample with certainty (i.e., probability = 1).
- 2) Cases not in the first group that were screened as businesses or non-working telephone numbers by GENESYS-IDplus. These cases were included in the subsample with certainty (i.e., probability = 1).
- 3) 10,000 cases not in either of the first two groups that were selected using the NIS outcomes as strata. These cases were included in the subsample with less than certain probability (i.e., probability < 1).

Appendix III

Computing Sampling Weights

When this report is edited, typeset, and printed, this appendix will present the methodology used to compute the sampling weight for the NSCH sample. The weighting procedure for the NSCH sample involves the following steps:

- Computing the base weight separately for each quarter
- Combining all quarters and computing full sample weights
- Adjustment for nonresolution of released telephone numbers by quarter
- Adjustment for incomplete age-eligibility screener by quarter
- Adjustment for subsampling of children within households
- Adjustment for incomplete NSCH interviews
- Adjustment for multiple telephone landlines
- Adjustment for noncoverage of nontelephone households
- Raking adjustment to external control totals.

Each of these weighting steps will be discussed in detail. This appendix will include [Table I](#) and [Table II](#).

Appendix IV

Questionnaire

The questionnaire for the 2007 NSCH (including question wording, response options, and instructions for skipping questions) will be added when this report is edited, typeset, and printed. Until such time, this questionnaire may be found on the SLAITS website (<http://www.cdc.gov/nchs/slaits.htm>).

Appendix V

Summary of Key Differences Between the 2003 and 2007 Questionnaires

Section 1

The question that asked the highest level of education attained by any member of the household (S1Q05A) was dropped from the 2007 version of the questionnaire. However, questions about the mother's, father's, and/or respondent's highest level of education can be found in Section 11 (K11Q20-22).

Section 2

The 2003 survey included a single question about specific dental problems (S2Q55 INDEX) that was asked only when parents reported that the child's teeth were in fair or poor condition. For 2007, new questions about specific dental problems (K2Q52-55) were asked for all children 12 months of age or older regardless of the reported condition of the child's teeth. Also for 2007, the question about the condition of the child's teeth (K2Q01_D) was moved earlier in the interview and away from the questions about specific problems.

For 2007, a new question was added to ascertain the child's birth weight (K2Q04). This question was asked only for children less than six years old because of concerns about the reliability of parent recall for older children.

The 2003 survey included questions that asked whether a doctor or other health professional had ever told the parent that the child had various chronic health conditions. For 2007, new follow-up questions asked parents to report whether the child still had each condition.

The 2003 survey grouped depression and anxiety problems (S2Q22) and hearing problems and vision problems (S2Q20). For 2007, questions asked about each of these conditions separately (K2Q32A, K2Q33A, K2Q43A, K2Q44A). In addition, for 2007, new questions were added to determine whether a child had Tourette Syndrome (K2Q38A) or epilepsy (K2Q42A).

The 2003 survey included a single question that assessed the overall severity of all conditions for all children who had ever been diagnosed with one or more of the chronic conditions included in the survey. For 2007, separate questions about severity were asked for each health problem that the child was reported to have.

Several questions in the 2003 survey about asthma and mental and emotional health were dropped for 2007. These questions included whether or not the child had emotional, concentration, or behavior difficulties (S2Q59); the degree to which the child's asthma (S2Q51) or mental and emotional health (S2Q61) placed a burden on the family; hospitalization due to asthma (S2Q53); asthma attacks (S2Q52A); health difficulties due to asthma (S2Q50); and asthma medication use (S2Q52).

For 2007, new questions on activity limitations (K2Q60A-K2Q61C) were asked for children with special health care needs.

The 2003 survey included questions about immunizations for Hepatitis A (S4Q27-S4Q30). These questions were dropped for 2007. For 2007, questions were added about immunizations for adolescents aged 12 to 17 years old (K2Q81-K2Q85). Questions were asked about receipt of the tetanus booster (Td or Tdap), meningitis shot (Menactra or Menomune), and the human papillomavirus shot (Gardasil, or Cervarix).

Section 3

The 2003 survey included a question about dental insurance coverage (S3Q03). This question was dropped for 2007.

For 2007, questions about the adequacy of the child's health insurance coverage were added (K3Q20-K3Q21B) for those children insured at the time of the interview.

Section 4

This section was substantially revised for 2007. The 2003 survey included questions that determined the number of times the child visited an emergency room (S4Q04), whether the child needed prescription medications (S4Q15-S4Q16), and why the child did not receive certain kinds of care (S4Q08 INDEX, S4Q14 INDEX, & S4Q18 INDEX). These questions were dropped for 2007.

For 2007, several new questions were added to ascertain whether the child has a usual place for care (K4Q01-K4Q03). Additional new questions asked about access and utilization of care from mental health professionals (K4Q22) and specialty doctors (K4Q24-K4Q28).

The 2003 survey included a question asking if the child received all needed routine preventive dental care (S4Q13). For 2007, this question was modified to ask how many times the child saw a dentist for preventive dental care (K4Q21).

For 2007, a question was added to ascertain whether children used any medication because of difficulties with emotions, concentration, or behavior (K4Q23).

Section 5

This section was substantially revised for 2007. Nearly all of the questions in the 2003 survey focused on the availability of, and the quality of care received from, the child's personal doctor or nurse; these questions (S5Q02-S5Q09, S5Q09B-S5Q10C) were dropped.

For 2007, this section included many of the same questions used to assess the medical home in the 2005-2006 National Survey of CSHCN. These questions determined the ease of obtaining needed referrals (K5Q10-K5Q11), the need for care coordination (K5Q20-K5Q22), the satisfaction with health care providers' communication (K5Q30-K5Q32), and the receipt of family-centered care (K5Q40-K5Q44).

Section 6

For 2007, new questions were added regarding the receipt of a formal developmental screening (K6Q12-K6Q14B) and whether the child had an Individualized Family Service Plan or Individualized Education Program (K6Q15).

The 2003 survey included questions about accidental poisoning (S6Q57-58 INDEX). These questions were dropped for 2007.

The 2003 survey included questions about five different types of child care (S6Q48-S6Q52). For 2007, these questions were dropped in favor of new questions that determined whether child care was received from relatives or nonrelatives and whether that care was provided in the child's home, another home-based setting, or a center (K6Q20-K6Q22).

For 2007, new questions were added for children with special health care needs to determine whether their health or behavior limits the parents' ability to find child care (K6Q25B-K6Q25C).

The 2003 survey asked all respondents about the need to make different arrangements for child care at the last minute (S6Q53). For 2007, this question was limited to children for whom child care was used or needed.

The 2003 survey included questions about the duration of breastfeeding. For 2007, additional questions were added about exclusive breastfeeding, including the age when the child was first fed formula (K6Q42) and the age when the child was first fed anything other than breast milk or formula (K6Q43).

The 2003 survey asked one question about reading stories to the young child. For 2007, a new question was added about telling stories and singing songs (K6Q61), to better address cultural variations in early literacy activities.

For 2007, new questions were added to determine how often young children play with other children their age (K6Q63) and how often young children watch television or videos (K6Q65).

Section 7

For 2007, new questions were added to determine whether the child had an Individualized Education Program (K7Q11).

The 2003 survey included a question regarding the number of days that the child participated in clubs, organizations, or sports teams (S7Q12). For 2007, this question was dropped.

For 2007, the question about self-care was modified to more clearly state that this question refers to time spent alone without an adult or teenager due to difficulties in making care arrangements (K7Q35).

For 2007, the response options for the question about volunteer work (K7Q37) were modified to ascertain the frequency of such work.

For 2007, the question about the number of hours that the child works for pay (K7Q39) was preceded by a new question asking whether the child earns money for work (K7Q38). This new question also clarified that occasional work should be included along with regular jobs.

The 2003 survey included questions about helmet use (S7Q22-S7Q23). For 2007, these questions were dropped.

For 2007, a new question was added to determine whether there is a television in the child's bedroom (K7Q62).

The 2003 survey included questions about the respondent's specific concerns regarding the child (S7Q30-S7Q39). For 2007, these questions were dropped.

For 2007, the answer choices for the NSCH Social Competence Scale (K7Q70-K7Q77) and for three questions about depressed mood (K7Q78-K7Q80) were modified to include "rarely" as an option.

For 2007, two new questions were added to determine the child's level of school engagement (K7Q82-K7Q83).

Section 8

The 2003 survey included questions about the degree of closeness in the respondent's relationship with the child (S8Q04) and about how members of the household deal with serious disagreements (S8Q12-15). For 2007, these questions were dropped.

For 2007, the answer choices for the Aggravation in Parenting Scale (K8Q31-K8Q34) were modified to include "rarely" as an option. Also, one of the aggravation questions included in the 2003 survey (S8Q09) was dropped for 2007 due to limited cultural validity among Spanish-speaking Latino parents. These parents were more likely to positively regard the need to give more to meet the child's needs than they had expected.

Section 9

For 2007, new questions were added to determine whether the respondent had legally adopted the child (for children not living with any biological parent; K9Q12_A), the age of the child's mother (K9Q16), the respondent's marital status (K9Q17A-K9Q17D-3) and marital happiness (K9Q18), and tobacco smoking inside the child's home (K9Q41).

The 2003 survey included a question for children not living with a biological parent that asked how often the child had seen his/her nonresident biological parent in the past twelve months (S9Q05-S9Q05A). For 2007, this question was dropped.

The 2003 survey included questions about the health insurance coverage of the child's parents (S9Q15C-S9Q15E). These questions were dropped for 2007.

Section 10

For 2007, new questions were added regarding neighborhood amenities (K10Q11-K10Q14) and the neighborhood condition (K10Q20-K10Q23).

The 2003 survey included questions about people in the neighborhood who might be a “bad influence” on the child (S10Q04) and about perceptions of the child’s safety at home (S10Q08). For 2007, these questions were dropped.

Section 11

For 2007, a new question was added for American Indian or Alaska Native children to determine whether services had been received services from any Indian Health Service hospital or clinic during the prior 12 months (K11Q03).

For 2007, three adoption-related questions were added to determine whether the child was adopted from another country (K11Q38) or from the U.S. foster care system (K11Q40), and whether the child’s adoption had been finalized (K11Q41).

The 2003 survey included a question about receipt of benefits from the Women, Infants, and Children (WIC) program (S9Q34). For 2007, this question was dropped.

For 2007, new questions were added to confirm the respondent’s ZIP Code (K11Q81-82) and state of residence (K11Q83).

Appendix VI

Summary of Questionnaire Changes during Data Collection

On May 2nd, 2007, changes were made to the following questions:

- K3Q04 was changed from “During the past 12 months, has he/she had health care coverage?” to “During the past 12 months, was there any time when he/she had health care coverage?”
- The skip patterns in questions K9Q17B_1 and K9Q17C_1 were adjusted to skip to K9Q18 if the respondent answered no, don’t know, or refused to these questions. The questions previously skipped to K9Q17B_2 and K9Q17C_2, respectively.
- The following questions were adjusted to allow answers from NIS to be filled when appropriate: K11Q02, K11Q02_OS, and K11Q20.
- Interviewer help text was added at K6Q25A to help clarify that “need for occasional babysitting” should not be included when reporting whether child care was needed.
- A system check was added at K9Q00 to ensure that the total number of people in the household was not equal to or less than the total number of children indicated in the household.

On July 5th, 2007, the following changes were made:

- A question for the interviewers was added at the end of the NSCH interview to ensure that non-English cases did not continue on to the NSAP interview.
- The text at K2Q45D was adapted from “Are (S.C.)’s current problems related to his/her bones, joints, or muscles?” to “Are (S.C.)’s current problems related to his/her bones, his/her joints, or his/her muscles?” to make the response options clearer for respondents.
- The skip logic at K2Q60A was revised to allow this question for children aged 12 to 71 months. Previously, the question was asked for all children under 6 years (72 months) of age.
- The skip logic at K2Q60C was revised to allow this question for children aged 36 to 71 months. Previously, the question was asked for all children under 6 years (72 months) of age.
- Text fills for K9Q11 were changed to add “other” where appropriate based on the relationship of the respondent to the child: “Does (S.C.) have any other parents, or other people who act as his/her parents, living here?”
- Wording at K9Q1159A was changed from “Would you say this income was above or below \$(dollar amount)?” to “Would you say this income was more or less than \$(dollar amount)?”
- The response option “U.S. territories” was added as an option at question K11Q83.
- The skip logic at K6Q11 was erroneously modified so that K6Q11 was asked only when parents reported concerns at K6Q01 and also reported a concern at K6Q02 through K6Q09. Previously, K6Q11 was asked for age-eligible children if any concerns were reported at K6Q01 through K6Q09. This error was not identified until after data collection had been completed.
- The text at K10Q22 was modified from “How about poorly kept or dilapidated housing?” to “How about poorly kept or rundown housing?” to aid respondents’ understanding of the question.

- K2Q38A was moved after K2Q41A (from before K2Q40A) to facilitate questionnaire flow.
- The system was adapted to allow answers to the health insurance questions from the NIS to be filled for the NSCH when appropriate (i.e., when the selected child in the NSCH was an NIS child).
- Six questions related to special health care needs were added for children who were not identified as having special health care needs. These included K2Q12A, K2Q15A, K2Q18A, K2Q21A, K2QTEST1, and K2QTEST2.
- K2Q30D was added.
- For questions K6Q65, K7Q50, K7Q51, and K7Q60, interviewer instructions (regarding “average weekdays recently”) were added to account for the transition from school year to summer activities.
- A system soft check was added at S_UNDR18 to verify entries less than S_NUMB (the NIS variable capturing how many NIS-eligible children were in the household).

On July 18th, 2007, the following changes were made:

- In SCQ02, SCQ04, and SL_INTRO, the sentences were rearranged so that the estimated interview length was provided after the respondent was offered a monetary incentive, if applicable.
- The response options for question K8Q23 were changed from “very often, sometimes, not very often, almost never, or never” to “never, rarely, sometimes, usually, or always.”
- The skip logic at K8Q26 was revised to allow this question for children aged 13 years or older. Previously, the question was asked for children aged 12 years or older.

On October 5th, 2007, the following changes were made:

- Interviewer instructions were added at K9Q20, K9Q21, K9Q22, K9Q23, K9Q24, K9Q25, K9Q30, K9Q31, K9Q32, K11Q20, K11Q21, K11Q22, K11Q30, K11Q31, K11Q32, K11Q34A, K11Q35A, and K11Q36A to clarify exactly who the question was referring to in cases where the respondent switched or a data entry error had occurred.
- The consent language was changed at SCQ02, SCQ04, and SL_INTRO from “You may choose not to answer any question you don’t wish to answer or stop at any time without penalty” to “You may choose not to answer any questions you don’t wish to answer, or end the interview at any time.” In addition, at SL_INTRO, the language was changed from “We will take all possible steps to protect your privacy and are required by law to use your answers only for statistical research” to “We are required by Federal law to develop and follow strict procedures to protect your information and use your answers only for statistical research.”

On January 3rd, 2008, the following changes were made:

- Language in SCQ02 and SCQ03 was adjusted to allow for smooth transitions from the NIS-Teen survey into the NSCH. In addition, the system was programmed to allow answers from the NIS-Teen survey to be filled for the NSCH when appropriate.

- The skip logic before K9Q11 was changed so that single-mother and single-father households skipped to K9Q12_A rather than K9Q16.

Appendix VII

Procedures for Assigning Poverty Status

The U.S. Department of Health and Human Services (DHHS) publishes Federal Poverty Guidelines for the determination of household poverty status (<http://aspe.hhs.gov/poverty>). These guidelines are produced annually and developed separately for three groups: the 48 contiguous states (plus the District of Columbia), Alaska, and Hawaii.

The 2007 NSCH used DHHS guidelines to assign household poverty status. Year 2007 guidelines (Tables III, IV, and V) were used with 2006 income for interviews conducted from April 5, 2007 through December 31, 2007, and with 2007 income for interviews conducted from January 1, 2008 through July 27, 2008. The tables were used to group households into the following poverty status categories:

- Category AA - At or below 50% of poverty level
- Category A - Above 50% to at or below 100% of poverty level
- Category B - Above 100% to at or below 133% of poverty level
- Category C - Above 133% to at or below 150% of poverty level
- Category D - Above 150% to at or below 185% of poverty level
- Category E - Above 185% to at or below 200% of poverty level
- Category F - Above 200% to at or below 300% of poverty level
- Category G - Above 300% to at or below 400% of poverty level
- Category H - Above 400% of poverty level

Two variables were used to determine household poverty status: the number of people residing in a household and the total household income during the prior year. It was possible for income data to be gathered using one of three different methods. A respondent could provide an exact income, provide an income range based on a closed-ended series of questions, or provide an income range using a set of cascade questions revised to allow exact determination of household poverty status in cases where that would not otherwise be possible. A brief description of each method and the household poverty status assignment process appears in the following text.

Respondent Reported Exact Income—When a respondent reported an exact income, poverty status was assigned by simply comparing the number of household members and the exact income reported with the appropriate guidelines table.

Respondent Reported Income Range Based on a Closed-Ended Series of Questions—When respondents did not supply a specific dollar amount for household income, it was necessary to ask a series of questions to determine whether the household income was below, exactly at, or above threshold amounts. A matrix (Tables VI, VII, and VIII) was created to categorize these responses. Each cell in the matrix was assigned to one of the following income categories:

- \$0 to \$7,500

- \$7,501 to \$10,000
- \$10,001 to \$12,500
- \$12,501 to \$15,000
- \$15,001 to \$17,500
- \$17,501 to \$20,000
- \$20,001 to \$25,000
- \$25,001 to \$30,000
- \$30,001 to \$35,000
- \$35,001 to \$40,000
- \$40,001 to \$45,000
- \$45,001 to \$50,000
- \$50,001 to \$60,000
- \$60,001 to \$75,000
- \$75,001 or higher

Respondents who went through the cascade of income questions were assigned a household poverty status by comparing the number of household members and the income range obtained through the income cascade with the appropriate poverty level guidelines table. When respondents did not complete the income cascade, either because they refused or did not know the answer to one of the cascade questions, household poverty status could not be assigned.

Respondent Reported Income Range Based on Revised Series of Cascade Questions—In some cases, the income categories described previously encompassed one or more income breaks for determining household poverty status. In such cases, additional income cascade questions beyond the standard set were asked to permit definitive assignment of poverty status. For example, the income break indicating that a two-person household in the contiguous 48 states was below 50% of poverty (using the 2007 guidelines) was \$6,845. This income break is encompassed in the income category of “\$0 to \$7,500.” Therefore, an additional cascade question asked whether the household income was more than, exactly, or less than \$6,800 (a reference value based on rounding rules described in the notes to the poverty guidelines tables). Based on the answers to these additional questions, the reported income ranges no longer encompassed poverty status breaks. Household poverty status was assigned by comparing the number of household members and the income range with the appropriate guidelines table.

Using DHHS guidelines and NSCH-specific rounding rules, tables were developed to provide reference values for the additional income cascade questions. These reference values are presented in [Tables VI, VII, and VIII](#).

Appendix VIII

Program Names Used for Medicaid and SCHIP Questions

For questions regarding Medicaid and the State Children’s Health Insurance Program (SCHIP), the state-specific program names for each type of coverage were included in the question text, in case respondents recognized the state program name but not the national program affiliation. These program names are shown in **Table IX**. Because a single question (K3Q02) was asked about both Medicaid and SCHIP, survey analysts will not be able to distinguish between Medicaid and SCHIP coverage in national or regional analyses. Analysts may be required to report on “public” insurance only.

Appendix IX

Letters Sent to Sampled Households

This appendix contains the full complement of advance letters, follow-up letters, and thank you letters used over the course of data collection. The following 9 letters are included in this appendix:

- 1) NIS advance letter for quarter 2, 2007. This letter was also used for some NIS households in quarters 3 and 4, 2007.
- 2) NIS advance letter for quarter 1, 2008. This letter was one of two experimental letters used for some NIS households in quarters 3 and 4, 2007.
- 3) NIS experimental advance letter for quarters 3 and 4, 2007.
- 4) Advance letter for banked sample (with mention of the NIS).
- 5) Advance letter for banked sample (without mention of the NIS).
- 6) Advance letter for NSCH-only “augmentation” households.
- 7) Follow-up letter when incentives were offered to households that had refused twice.
- 8) Follow-up letter when incentives were offered to households that could not be successfully contacted over a period of time.
- 9) Thank you letter when incentive was mailed. The amount of the incentive listed was either \$10 or \$15, depending on whether the household already received \$5 in an advance letter.

Because the NSCH typically follows the NIS, the advance letter sent to most households was the usual NIS advance letter. In the second half of 2007, the NIS studied whether response rates were affected when the advance letter was altered to improve readability and respondent comprehension. (The primary content of the advance letter remained unchanged.) The experiment utilized three letters in Q3/2007 and Q4/2007: the original NIS letter that was used in Q2/2007 and two experimental letters. The first experimental letter included changes to the overall tone and formatting, as well as placement of confidentiality content based on recommendations of the U.S. Census Bureau’s 2001 report on the Survey of Income and Program Participation (SIPP) advance letter (31). The second experimental letter provided short paragraphs about the study on the front, followed by labeled sections to address common questions and provide full information about the study on the back. To evaluate the experiment, the NIS considered age-eligibility rates, cooperation rates, and response rates. None of the three letters consistently produced the highest rates. The NIS adopted the first experimental letter as the NIS advance letter beginning in Q1/2008.

For the Q1/2007 banked sample, significant time had elapsed since cases received the NIS advance letter. Therefore, an NSCH-specific advance letter was sent to all cases with an available address. An experiment comparing two different letters was conducted to determine whether referencing the NIS and that survey’s previous call attempts would trigger a memory for the respondent and whether that would be an asset or a detriment to the NSCH. The letter mentioning the NIS included the following text as the first paragraph:

Recently, the Centers for Disease Control and Prevention (CDC) asked your household to take part in the National Immunization Survey (NIS). Thank you for the time you took to respond to

that survey. We now need your help with another important health survey about children across the U.S.

Beyond this language, the letters had no significant differences. No significant difference in response rates between the two letters was found.

When this report is edited, typeset, and printed, the letters will be added.

Appendix X

Disposition Code Frequencies and Response Rate Calculation

This appendix consists of **Table X** and **Table XI**.

Appendix XI

Incentive Effort

To improve the likelihood that age-eligible households would participate in the NSCH and would contribute to a more complete dataset, an incentive plan was developed and executed during NSCH data collection. This plan was guided by previously successful incentive efforts in the NIS and the 2005-2006 National Survey of CSHCN. Although initial consideration was given to offering an incentive payment to all age-eligible NSCH households who completed the interview, cost constraints prohibited the application of this model in light of the significant number of targeted interviews planned for the NSCH.

Eligibility

A two-tiered incentive model was constructed to offer incentives to age-eligible households that refused participation in the NSCH. The first refusal could have occurred during NIS, NIS-Teen, or NSCH. Once a case refused, it became eligible for Tier 1 of the incentive model. A \$10 incentive for NSCH interview participation was offered to the household during the next call attempt. If the case refused a second time without completing the interview, the case became eligible for Tier 2 of the incentive model. An additional \$5 was offered to the household. Households that refused after this additional incentive offer without completing the interview were finalized (i.e., given a final disposition code) and not called again.

Eligibility for an NSCH incentive was initially limited to age-eligible households that had refused participation (i.e., an active refusal). However, nonresponse analyses of age-eligible households following implementation of the two-tiered incentive model led to the application of an expanded incentive effort. This analysis of age-eligible cases that had neither finalized nor finished the NSCH interview revealed two groups of cases with similar patterns of nonresponse. The first group consisted of households that had not refused participation, but multiple attempts to contact these households over 21 or more days had resulted in no contact. The second group consisted of households that had refused participation once and had become eligible for the Tier 1 incentive, yet multiple attempts to recontact these households over 21 or more days had resulted in no contact. These patterns of continued nonresponse despite repeated attempts at contact with the household were classified as passive refusals. The Tier 2 incentive (\$15), typically reserved for active refusal cases, was offered to passive refusal cases to ensure they were not underrepresented in the NSCH data and to maximize the likelihood of contact with these households.

Following the passive incentive offer, both types of passive refusal cases remained active until they accumulated two refusals in their call history. Passive refusal cases with one previous active refusal were finalized after one subsequent refusal; passive refusals with no previous active refusals were finalized after the second subsequent refusal. At the point of their second refusal, both groups of passive incentive cases were finalized and were not called again.

If, at any time, a household refused participation in a hostile manner or requested to be removed from the calling list, the case was deactivated and not called again. The case was not eligible for Tier 1 or Tier 2 incentives.

Procedure for Offering Tier 1 Incentives

Once a household met the criteria for the Tier 1 incentive, the incentive offer of \$10 was introduced in the consent script, callback script, and/or answering machine script, depending on the point at which the refusal occurred. For households that were ineligible for the NIS interview and had refused prior to the informed consent process, the following script was read by the interviewer:

Before we continue, I'd like you to know that taking part in this research is voluntary. You may choose not to answer any questions you don't wish to answer, or end the interview at any time. We are required by Federal law to develop and follow strict procedures to protect your information and use your answers only for statistical research. I can describe these laws if you wish. In appreciation for your time, we will send you \$10. The survey will take about 25 minutes. In order to review my work, my supervisor may record and listen as I ask the questions. I'd like to continue now unless you have any questions.

Prior to July 18, 2007, the monetary incentive payment was offered to respondents after informing them of the expected interview duration. Beginning on July 18, 2007, the monetary incentive was offered immediately before informing respondents of the interview duration (as shown in the previous script). Consent scripts were tailored for cases that had already completed the NIS or NIS-Teen and had heard the NIS or NIS-Teen informed consent script.

For cases that had refused after completing part of the NSCH interview, the interviewers read the following script:

Hello, my name is _____. I'm calling on behalf of the Centers for Disease Control and Prevention. Earlier, someone in your household started an interview about the health of children and teenagers, and we began talking about one child in your household. I'm calling back now to continue the interview. In appreciation for your time, we will send you \$10.

Information about the incentive also was left as part of the answering machine message:

Hello. I am calling on behalf of the Centers for Disease Control and Prevention regarding a nationwide study about the health of children and teenagers. When we spoke previously about this important study, you requested that we call you back at this time. I'm sorry that we've missed you. We'll try to contact you again soon but please feel free to return our call anytime at 1-866-999-3340. In appreciation for your time, we will send you \$10. If you have any questions, that number again is 1-866-999-3340. Thank you.

If a household completed the NSCH interview, or if a respondent requested the incentive without completing the interview, address information for the household was either confirmed or collected. The \$10 cash payment was mailed to the household, along with a letter ([Appendix IX](#)) expressing appreciation for the respondent's time and effort spent participating in the interview. Incentive-eligible households that completed the NSCH interview, but declined to confirm or provide address information, were not mailed the incentive payment.

Procedure for Offering Tier 2 Incentives

Households that met criteria for the Tier 2 active refusal incentive or the passive refusal incentive were approached in a different manner. Upon delivering their second refusal or meeting eligibility criteria for either passive refusal group, cases were temporarily suspended within the CATI system.

For households with an available address, a letter ([Appendix IX](#)) was mailed with \$5 cash enclosed. The letter explained that attempts had been made to contact the household via telephone to complete the NSCH interview. It also briefly described the NSCH, included an FAQ section on the survey, and mentioned that \$10 would be mailed upon continued participation in the NSCH. Approximately two to four weeks after temporary suspension, the cases eligible for the Tier 2 incentive were reactivated. Households that had received the \$5 incentive in the mail were offered \$10 by telephone; households without an available address that had not received the \$5 incentive in the mail were offered the total \$15 incentive by telephone. The incentive offer was mentioned by the interviewer, using scripts similar to those used to offer the Tier 1 incentive.

If any Tier 2 incentive-eligible household completed the NSCH interview, or if a respondent requested the incentive without completing the interview, address information for the household was either confirmed or collected. A final cash payment of \$10 was mailed to households that had already received \$5 in the mail, and a final cash payment of \$15 was mailed to households that had not yet received any incentive by mail. In addition, a letter ([Appendix IX](#)) expressing appreciation for the respondent's participation accompanied each final incentive payment. Incentive-eligible households that completed the NSCH interview, but declined to confirm or provide address information, were not mailed the incentive payment.

Interview Completion Rates

The incentive offers for active refusals (Tier 1 and Tier 2) were first implemented on May 15, 2007. The expanded incentive effort for passive refusals (Tier 2) began on August 2, 2007. Active and passive refusals from all quarters of data collection (including the Q1/2007 banked sample) were eligible.

A total of 66,365 cases became eligible for some type of an incentive during the NSCH. Nearly one-half (47.9%) of these cases completed the interview subsequent to incentive eligibility. Information on interview completion rates by incentive type is displayed in [Table XII](#).

In [Table XII](#), note that any case with one refusal that did not respond to the \$10 incentive in Tier 1 could become eligible for Tier 2 as either an active or passive refusal. The 19,087 completed cases in Tier 1 finished the interview after the \$10 incentive was offered and were not eligible for additional incentives. Of the 40,865 eligible Tier 1 cases that did not complete the interview after the \$10 incentive was offered, 32,885 were offered the \$15 incentive: 28,297 as active refusals and 4,588 as passive refusals (with one active refusal). The remaining 1,824 cases classified as active refusals in Tier 2 were eligible because they refused twice during the NIS and/or NIS-Teen interview.

The expanded incentive effort for passive refusals was more successful in achieving completed interviews, relative to the two-tier incentive effort for active refusals. However, this difference was largely the result of considerable success completing interviews with households that had never refused

participation but had been very difficult to contact. Passive refusal incentive-eligible cases may be less resistant to completing the survey than cases that more actively refuse participation.

Table XIII presents completion rates by quarter of data collection for each incentive type. From Q2/2007 to Q1/2008, the completion rates improved for each incentive type. The high rates for the Q1/2007 banked sample may be attributed, in part, to the extended “cooling off” period between the NIS screening or interview and the start of NSCH interviewing.

Appendix XII

Multiple Imputation of Household Poverty Level

The 2007 NSCH provides a rich source of data for studying the relationships between income and health and for monitoring health, health care, and well-being for children at different income levels. However, as is common for most household interview surveys, nonresponse rates were high for the question on total combined household income for the previous calendar year. Answers to this question, along with answers to a question about the number of people living in the household, are used to create an index of income relative to the Department of Health and Human Services Federal Poverty Guidelines. If data for either of these two components were missing, refused, or had a “don’t know” response, the household poverty status indicator was assigned a missing value code in the publicly released datasets. (Further details about the procedures for assigning household poverty status are available in [Appendix VII](#).)

For the 2007 NSCH, poverty status is missing for 8.5% of the households (7,817 of 91,642 households). Missing values for poverty status were predominately the result of missing data for income rather than missing data for household size. Only 394 households did not report household size.

A nonresponse analysis shows that missingness is related to several variables, including items pertaining to health, neighborhood and community characteristics, and demographics. Thus, the respondents cannot be treated as a random subset of the original sample. It follows that the most common method for handling missing data in software packages, “complete-case analysis” (also known as “listwise deletion”), will generally be biased because this method deletes cases that are missing any of the variables involved in the analysis. Moreover, since deletion of incomplete cases discards some of the observed data, complete-case analysis is generally inefficient as well; that is, it produces inferences that are less precise than those produced by methods that use all of the observed data.

Imputation is a more appropriate approach to handling nonresponse on items in a survey for several reasons. First, imputation adjusts for observed differences between item nonrespondents and item respondents; such an adjustment is generally not made by complete-case analysis. Second, imputation results in a completed data set, so that the data can be analyzed using standard software packages without discarding any observed values. Third, when a data set is being produced for analysis by the public, imputation by the data producer allows the incorporation of specialized knowledge about the reasons for missing data in the imputation procedure, including confidential information that cannot be released to the public. Moreover, the nonresponse problem is addressed in the same way for all users, so that analyses will be consistent across users.

Although single imputation, that is, imputing one value for each missing datum, enjoys the positive attributes just mentioned, analysis of a singly imputed data set using standard software generally fails to reflect the uncertainty stemming from the fact that the imputed values are plausible replacements for the missing values but are not the true values themselves. As a result, analyses of singly imputed data tend to produce estimated standard errors that are too small, confidence intervals that are too narrow, and significance tests that reject the null hypothesis too often when it is true.

Multiple imputation is a technique that seeks to retain the advantages of single imputation while also allowing the uncertainty due to imputation to be reflected in the analysis (32). The idea is to first simulate $M > 1$ plausible sets of replacements for the missing values, which are then combined with the nonmissing values to generate M complete data sets. The M complete data sets are then analyzed separately using a standard method for analyzing complete data, and finally the results of the M analyses are combined in a way that reflects the uncertainty due to imputation. For public-use data, M is not usually larger than five, which is the value that has been used here in multiply imputing missing data for the NSCH.

This appendix describes the procedures used in multiply imputing household income and household size for the NSCH. Household poverty status is expressed as a percentage; households with income less than 100% of the federal poverty level are considered to be living in poverty. For each of the multiply imputed data sets, household poverty status was derived from the imputed values for household income and household size.

Imputation Procedures

Income and household size were each imputed five times, creating five imputed data sets. The literature on multiple imputation suggests that this is a sufficient number of imputations unless the amount of missing information is extreme (32). As noted earlier, the number of survey records with missing household size values was much smaller than the number of survey records with missing household income values. Because there was very little missingness in household size to explain, predictors for household size were not explored separately from predictors for household income. Therefore, household size was imputed using the same predictors used for household income. When both household size and household income were missing for a single case, five *pairs* of imputed values were produced.

The imputation of household income and household size was complicated by two issues. First, neither household income nor household size was normally distributed. This is a disadvantage because linear regression modeling assumes that the dependent variable being modeled has a normal distribution. Therefore, we used transformed variables for modeling and imputation. To determine the suitable transformation for income and household size to conform to the normality assumption in the imputation model, Box-Cox transformations were estimated from the observed data. For household size, the log transformation led to normality. For income, the optimal transformation was to the 0.22 power, which was rounded to the quarter-root (0.25). After the imputation procedure was completed, the imputed values were transformed back to their original scale.

Second, in some cases, the imputed values of household income and household size needed to be constrained within certain bounds. Household respondents were asked to provide an exact household income. However, when respondents did not provide an exact household income, a series (i.e., cascade) of questions asking whether the household income was below, exactly at, or above threshold amounts were then asked. The multiple imputation procedures employed for the NSCH needed to impute the income value so that it was consistent with any information gathered from the cascade questions. For households with missing data on household size, the imputed values needed to be restricted so that they

were consistent with other information provided in the survey (e.g., household size is greater than the number of children in the household).

The software IVEware, available online at <http://www.isr.umich.edu/src/smp/ive>, allows the user to specify lower and upper limits of imputed values, constraining the imputation distribution from which draws are made. This software has been used to impute family income and family earnings for the National Health Interview Survey and to impute household income and household size (to derive household poverty status) for the 2001 and 2005-2006 National Surveys of CSHCN and the 2003 NSCH.

IVEware uses sequential regression multivariate imputation (SRMI). With sequential regression imputations, income and household size had separate models that used the same covariates, including each other. This technique was not as robust as some other imputation techniques that specify a joint model for both income and household size conditional on the predictor variables (33, 34). However, this slight disadvantage of using SRMI is outweighed by IVEware's ability to constrain the imputed values within specified lower and upper limits.

IVEware builds regression models, and then multiply imputes variables based on the models built. For understanding relationships between variables, parsimony is desired, but in prediction (imputation can be thought of as "predicting" the missing values), more complicated models are often better for two reasons. First, using more variables leads to a higher correlation between the observed and predicted values for a model. Second, the validity of analyses conducted on multiply-imputed datasets is broader when more variables are included in the model.

In the imputation model, as many predictors as possible were included. To produce high-quality imputations, variables that were potentially related to household income and potentially related to the missingness of household income were included. Another important consideration was to include variables that account for features of the sampling design. Inclusion of variables to reflect the sampling design was necessary so that approximately valid inferences will be obtained when the multiply imputed data are analyzed.

The imputation model included variables related to the questionnaire items on demographics (for the child and family), health and functional status of the child, health insurance coverage, health care access and utilization, medical home, and neighborhood characteristics. Variables related to early childhood and middle childhood were not included, as these items, being targeted to a subset of the sample, had large percentages of missing values. For most of the variables, the "refused" or "don't know" answers were recoded to missing. For some variables having logical skips, logical imputation was used to obtain more complete variables. For example, the variable K11Q60 (receipt of cash assistance) is missing when the household's income does not qualify for the cash assistance. Therefore, it was recoded as a "no" response for such households. Also, many categorical variables were recoded with top-coding or bottom-coding to reduce the number of rarer categories. For example, for the variable K4Q20 (number of doctor visits), the values range from 0 to 365 with small frequencies for values greater than 10. The number of categories was reduced to 11 with category "10" defined as 10 or more visits. Another consideration used while building the imputation model was to not include covariates having more than 5% missing values.

Because fitting the regressions in the SRMI procedure does not automatically account for features of the sample design, variables reflecting the design were included as predictors in the regression models. The strata for this design were the 50 states and the District of Columbia. To account for the stratum effect, states, in the form of 50 indicator variables, and state-level income summary variables (mean and standard deviation with log transformation) were considered as possible covariates in the imputation model. The state indicator variables were dropped before the final imputations were carried out, but retained the state-level income summary variables. Survey weights were also included as covariates in the model, after transforming the weights to a logarithmic scale.

Results of Modeling

Table XIV shows the 47 variables (other than survey weight) chosen for the model by stepwise regression. Most, but not all, have a significant relationship with income. Those with negative parameters decrease predicted income, whereas those with positive parameters increase predicted income. The model was highly significant, $F(47, 70105) = 1,113.50$, $p < .0001$. The R-squared and adjusted R-squared values for this model are both 0.43.

It should be noted that the imputed values for family income were not obtained from this regression model. The imputed values were drawn from the posterior distribution of missing family income based on the model derived from this regression.

Table XIV also shows which of the 47 variables (except for the 2 continuous state-level income variables) were significant predictors of nonresponse. These results are based on odds ratios derived after fitting a logistic regression predicting whether income was missing. Odds ratios greater than 1 indicate that an affirmative response on the variable is associated with higher rates of missingness. Just over half (24 out of 45) of the variables were significant predictors of nonresponse.

Use of Multiply Imputed Values

The derived poverty level variable that is available for public use was calculated from household income and household size. When either or both were missing, the derived poverty level was calculated from the imputed values. Regardless of whether the derived poverty level was based on reported or imputed values, the variable has been given the same name (POVLEVEL_I). A flag (FLAG_I) indicates whether the derived poverty level was based on reported or imputed values.

When missing, household income and household size were imputed five times. Therefore, the resulting data set contains five times as many observations as were in the original data set. For the 2007 NSCH, the datasets have $5(91,642) = 458,210$ records. Each complete set of derived poverty level values is distinguished by the SAS variable IMPUTATION. Therefore, each IDNUMR appears five times in the file, with IMPUTATION having values of 1, 2, 3, 4, and 5 corresponding to the five separate complete sets of derived poverty level values.

The public use data files for the NSCH do not include household income, to protect against inadvertent disclosure of survey subjects' identities. Only poverty level is reported on the public use data files. Similarly, imputed household income will not be released as public use data. Researchers

interested in accessing the original and imputed household income data may access the data through the NCHS Research Data Centers.

There are three possible ways to analyze the data. One invalid way to use the data that should not be attempted is also described.

Taking the possible ways first, a complete-case (only) analysis is the simplest, which uses only the cases with observed values. This can be done by using the poverty level variable (POVERTY_LEVELR) in the NSCH data file. Any analysis using this variable could be biased due to nonresponse, and the variability will be larger because of the missing values.

The second possible way of using the data is to use only a single imputation from the multiple imputation files. Each of the five imputations has been drawn from a valid distribution based on a regression model, but this model and the distribution are slightly different for each imputation. To analyze only one imputation, choose only the subset of cases with IMPUTATION = c, where c is 1, 2, 3, 4, or 5. Single imputation analyses result in estimated standard errors that are too small because the imputed values are treated as if they were observed. This ignores the inherent uncertainty resulting from lack of knowledge about the true (unobserved) value, but is superior to the complete-case analysis. It should be noted that slightly different results will be obtained depending on which subset of cases is chosen, but no subset is superior to another.

The statistically valid way to analyze the data is to analyze all five imputed datasets together. To do this, five separate analyses are conducted; one on each of the five imputed datasets. These analyses are then combined following the standard multiple imputation combining rules (32). This is superior to the previous two methods.

It is very important to note that it is invalid to combine the five imputed values into one analysis. For example, taking the average poverty level (which might not be an integer) to derive one “average” poverty status value per case is invalid. Poverty status must be analyzed as a multiply imputed variable with SAS, SUDAAN, IVEware, or another appropriate statistical software package to make use of the multiply-imputed data.

Regardless of the statistical software used to analyze the data, one must merge the survey data from the public use analysis files with the data from the multiple imputation file by the unique household identifier (IDNUMR). To combine these files, first sort by IDNUMR and then merge using this identifier as the merge variable.

For SAS analyses, it is also very important to have the dataset sorted by IMPUTATION because analyses of the multiply imputed data need to be done separately by IMPUTATION. Separate analyses are specified in SAS by using the procedure option keyword BY (“BY IMPUTATION;” should be one line within the analysis). In SAS, the two basic steps to using the multiply-imputed data are to 1) analyze the data separately by IMPUTATION as if each were a separate data set, and 2) combine the results from the different imputed data sets using PROC MIANALYZE. In the first step, separate analyses are done with options set to keep the covariances that are needed to combine the analyses. Then, PROC MIANALYZE combines these different analyses using the standard multiple imputation combining rules.

For SUDAAN analyses, a separate analytical file is necessary for each of the five imputations. The five datasets should then be sorted by the stratum (STATE) and the primary sampling unit (IDNUMR) variables. To analyze the data using the five imputation files, the MI_COUNT command should be added to the SUDAAN procedure call. The MI_COUNT command tells SUDAAN how many imputation files to expect.

Further instructions and examples for using SAS and SUDAAN are available in the User's Guide included as part of an earlier report on the multiple imputation of missing household poverty level values from the 2003 NSCH (35). This report is available online.

Appendix XIII

Prevalence Estimates and Weighted Frequencies

This appendix consists of **Table XV**.

Table A. Percentage of
National Survey of
Children's Health sample
called only for the
National Survey of
Children's Health, by
state

State	Percent
Connecticut	17.8
Delaware	4.9
Idaho	18.8
Kansas	7.9
Mississippi	2.6
Montana	8.5
North Dakota	10.8
Oklahoma	17.0
Utah	8.0

NOTE: Percent for all
other states and the
District of Columbia was
zero.

Table B. External (nongovernmental) technical expert panel members

Name	Affiliation (in 2006)
Maja Altarac, M.D., Ph.D.	University of Alabama at Birmingham
Christina Bethell, Ph.D., M.B.A., M.P.H.	Oregon Health and Science University
Neal Halfon, M.D.	University of California, Los Angeles
William Hollinshead, M.D.	Rhode Island Department of Health
Charles Irwin, M.D.	University of California, San Francisco
Jeffrey Lobas, M.D., M.P.A.	Iowa Child Health Specialty Clinics
Kristin Anderson Moore, Ph.D.	Child Trends
Paul Newacheck, Dr.P.H., M.P.P. (chairperson)	University of California, San Francisco
Lynn Olson, Ph.D.	American Academy of Pediatrics
Edward Schor, M.D.	The Commonwealth Fund
Judith Shaw, Ed.D., M.P.H., R.N.	University of Vermont

Table C. Number of interviewers trained and certified by month and telephone center location

Month and year	Location					
	Chicago, IL		Las Vegas, NV		Both locations	
	Number trained	Number certified	Number trained	Number certified	Number trained	Number certified
2007						
March	184	183	103	101	287	284
April	162	153	37	36	199	189
May	51	50	0	0	51	50
June	0	0	16	14	16	14
July	37	33	22	18	59	51
August	86	82	28	25	114	107
September	38	34	22	19	60	53
October	4	4	0	0	4	4
November	0	0	0	0	0	0
December	0	0	0	0	0	0
2008						
January	47	45	0	0	47	45
February	85	84	85	84
March	21	21	21	21
Total	715	689	228	213	943	902

... Not applicable. Interviewing at the Las Vegas location ended on January 8, 2008.

Table D. Number of interviews by state

State	Completed	Partially completed	Total
National	90,557	1,085	91,642
Alabama	1,737	24	1,761
Alaska	1,714	25	1,739
Arizona	1,741	28	1,769
Arkansas	1,751	14	1,765
California	1,728	23	1,751
Colorado	1,780	21	1,801
Connecticut	1,871	18	1,889
Delaware	1,786	19	1,805
District of Columbia	1,779	22	1,801
Florida	1,777	20	1,797
Georgia	1,766	16	1,782
Hawaii	1,780	42	1,822
Idaho	1,755	13	1,768
Illinois	1,897	35	1,932
Indiana	1,749	15	1,764
Iowa	1,744	13	1,757
Kansas	1,784	19	1,803
Kentucky	1,782	21	1,803
Louisiana	1,832	36	1,868
Maine	1,735	17	1,752
Maryland	1,755	18	1,773
Massachusetts	1,760	26	1,786
Michigan	1,847	14	1,861
Minnesota	1,753	14	1,767
Mississippi	1,880	38	1,918
Missouri	1,830	17	1,847
Montana	1,768	20	1,788
Nebraska	1,806	21	1,827
Nevada	1,796	25	1,821
New Hampshire	1,758	15	1,773
New Jersey	1,796	31	1,827
New Mexico	1,796	30	1,826
New York	1,751	42	1,793
North Carolina	1,758	24	1,782
North Dakota	1,784	16	1,800
Ohio	1,746	19	1,765
Oklahoma	1,797	19	1,816
Oregon	1,783	11	1,794
Pennsylvania	1,744	23	1,767
Rhode Island	1,740	16	1,756
South Carolina	1,857	28	1,885
South Dakota	1,724	16	1,740
Tennessee	1,813	27	1,840
Texas	1,783	22	1,805
Utah	1,741	6	1,747
Vermont	1,711	14	1,725
Virginia	1,755	19	1,774
Washington	1,721	15	1,736
West Virginia	1,746	20	1,766
Wisconsin	1,841	19	1,860
Wyoming	1,729	19	1,748

Table E. Number of interviews by month

Month	Total ¹	Percent
April 2007	4,435	4.8
May 2007	9,074	9.9
June 2007	9,449	10.3
July 2007	10,538	11.5
August 2007	11,208	12.2
September 2007	8,679	9.5
October 2007	7,454	8.1
November 2007	6,137	6.7
December 2007	5,269	5.8
January 2008	5,472	6.0
February 2008	4,542	5.0
March 2008	5,331	5.8
April 2008	2,533	2.8
May 2008	849	0.9
June 2008	460	0.5
July 2008	212	0.2
All months	91,642	100.0

¹Total number of interviews includes partially completed interviews.

Table F. Number and percentage of respondents by relationship to sampled child

Relationship	Number	Percent
Total	91,642	100.0
Mother	67,388	73.5
Father	18,759	20.5
Grandparent	3,854	4.2
Aunt or uncle	609	0.7
Sister or brother	531	0.6
Other guardian	483	0.5
Don't know/refused	18	0.0

0.0 Quantity more than zero but less than 0.05.

Table G. Mean and median length of National Survey of Children's Health interview in minutes and seconds, by section and by National Immunization Survey eligibility

Section of interview	NIS-eligible or NIS-Teen-eligible households		NIS-ineligible and NIS-Teen-ineligible households	
	Mean	Median	Mean	Median
Overall	22:55	21:54	28:05	26:56
Screeners: Age Eligibility, Selection of Sampled Child and Informed Consent	0:58	0:47	1:39	1:28
Section 1: Initial Demographics	0:20	0:14	0:34	0:29
Section 2: Health and Functional Status	4:43	4:23	5:21	5:00
Section 3: Health Insurance Coverage	1:04	1:01	1:23	1:19
Section 4: Health Care Access and Utilization	2:13	2:06	2:20	2:12
Section 5: Medical Home	2:01	1:54	2:15	2:09
Section 6: Early Childhood (0-5 years)	4:38	4:24	5:20	5:06
Section 7: Middle Childhood and Adolescence (6-17 years)	5:42	5:20	5:54	5:32
Section 8: Family Functioning	1:32	1:24	1:43	1:35
Section 9: Parental Health	2:09	1:58	2:17	2:07
Section 10: Neighborhood and Community Characteristics	1:57	1:49	2:03	1:54
Section 11: Additional Demographics	1:26	1:18	2:56	2:42

NOTES: NIS is National Immunization Survey. NIS eligibility refers to household eligibility. NIS-eligible households include at least one child between 19 and 35 months of age; households eligible for the NIS-teen interview include at least one child between the ages of 13 and 17 years. The NIS-eligible child in the household may or may not have been the child sampled for the National Survey of Children's Health.

Table H. Final disposition of the survey sample

Final disposition	Number of selected telephone lines	Percent of total selected telephone lines
Total	2,806,416	100.0
Not resolved as residential or nonresidential	445,972	15.9
Out of scope (i.e., business, nonworking, fax or modem)	1,770,887	63.1
Known household, age eligibility not determined	74,051	2.6
Age-screened household, no child in age range	380,130	13.5
Known age-eligible household, interview not completed	43,734	1.6
Known age-eligible household, partially completed interview	1,085	0.0
Known age-eligible household, completed interview	90,557	3.2

0.0 Quantity greater than zero but less than 0.05.

Table J. Weighted response rates, nationally and by state

State	Resolution rate	Age-screener completion rate	Interview completion rate	Overall response rate ¹
National	81.9	86.4	66.0	46.7
Alabama	82.9	87.2	70.1	50.6
Alaska	85.4	88.9	68.8	52.2
Arizona	80.5	85.4	64.3	44.2
Arkansas	86.1	89.7	69.7	53.8
California	77.2	83.4	62.2	40.1
Colorado	83.3	87.8	67.3	49.2
Connecticut	79.2	84.6	65.0	43.5
Delaware	78.0	86.0	65.4	43.9
District of Columbia	80.4	86.6	67.8	47.2
Florida	80.5	85.5	61.4	42.2
Georgia	83.2	85.4	63.0	44.7
Hawaii	85.0	84.2	59.0	42.2
Idaho	84.5	89.1	71.9	54.1
Illinois	83.7	86.4	64.7	46.8
Indiana	85.6	88.8	69.9	53.1
Iowa	86.8	90.3	72.5	56.8
Kansas	85.9	89.3	69.9	53.7
Kentucky	84.5	88.5	65.9	49.2
Louisiana	86.1	85.6	67.9	50.0
Maine	82.6	88.9	71.5	52.5
Maryland	79.3	85.1	65.7	44.3
Massachusetts	78.5	84.8	67.3	44.8
Michigan	84.3	88.5	67.4	50.3
Minnesota	85.0	89.1	69.3	52.5
Mississippi	84.5	86.7	65.8	48.2
Missouri	85.7	89.4	67.7	51.9
Montana	87.3	89.6	73.2	57.2
Nebraska	88.1	89.5	74.4	58.6
Nevada	78.5	84.3	63.1	41.7
New Hampshire	79.5	87.2	67.8	47.0
New Jersey	77.7	83.8	60.5	39.4
New Mexico	83.8	86.6	68.4	49.6
New York	81.0	84.4	60.7	41.5
North Carolina	81.5	87.7	68.2	48.7
North Dakota	89.2	90.6	76.6	61.9
Ohio	84.3	88.1	68.1	50.6
Oklahoma	83.5	87.5	66.9	48.9
Oregon	85.0	88.9	66.6	50.3
Pennsylvania	81.1	87.8	68.7	48.9
Rhode Island	80.8	86.1	69.2	48.2
South Carolina	81.8	86.9	66.6	47.3

South Dakota	88.8	90.5	73.2	58.8
Tennessee	83.3	87.6	70.6	51.5
Texas	81.3	84.8	66.6	45.9
Utah	86.1	88.1	71.8	54.4
Vermont	82.7	89.9	72.2	53.6
Virginia	80.6	87.4	67.5	47.5
Washington	83.6	87.0	66.4	48.3
West Virginia	80.1	88.3	66.8	47.3
Wisconsin	84.0	89.7	70.8	53.4
Wyoming	82.6	89.5	73.8	54.5

¹The overall response rate is the product of the resolution rate, the age-screener completion rate, and the interview completion rate.

Table K. Alternative weighted response rates, nationally and by state

State	Resolution rate	Age-screener completion rate	Interview completion rate	Overall response rate
National	89.9	86.4	66.0	51.2
Alabama	89.0	87.2	70.1	54.4
Alaska	94.1	88.9	68.8	57.5
Arizona	89.4	85.4	64.3	49.1
Arkansas	92.0	89.7	69.7	57.5
California	88.2	83.4	62.2	45.8
Colorado	90.7	87.8	67.3	53.6
Connecticut	87.3	84.6	65.0	48.0
Delaware	87.9	86.0	65.4	49.5
District of Columbia	93.4	86.6	67.8	54.9
Florida	88.4	85.5	61.4	46.4
Georgia	90.3	85.4	63.0	48.6
Hawaii	91.3	84.2	59.0	45.4
Idaho	91.6	89.1	71.9	58.7
Illinois	91.8	86.4	64.7	51.3
Indiana	92.1	88.8	69.9	57.2
Iowa	91.9	90.3	72.5	60.1
Kansas	92.2	89.3	69.9	57.6
Kentucky	90.7	88.5	65.9	52.9
Louisiana	92.7	85.6	67.9	53.9
Maine	89.3	88.9	71.5	56.7
Maryland	89.2	85.1	65.7	49.9
Massachusetts	86.4	84.8	67.3	49.4
Michigan	91.5	88.5	67.4	54.5
Minnesota	91.4	89.1	69.3	56.5
Mississippi	91.4	86.7	65.8	52.1
Missouri	92.0	89.4	67.7	55.7
Montana	92.5	89.6	73.2	60.6
Nebraska	92.8	89.5	74.4	61.7
Nevada	87.6	84.3	63.1	46.6
New Hampshire	86.6	87.2	67.8	51.2
New Jersey	88.6	83.8	60.5	44.9
New Mexico	90.2	86.6	68.4	53.4
New York	89.6	84.4	60.7	45.9
North Carolina	89.5	87.7	68.2	53.5
North Dakota	93.4	90.6	76.6	64.9
Ohio	89.8	88.1	68.1	53.9
Oklahoma	90.9	87.5	66.9	53.3
Oregon	91.3	88.9	66.6	54.1
Pennsylvania	89.4	87.8	68.7	54.0
Rhode Island	86.7	86.1	69.2	51.7
South Carolina	89.5	86.9	66.6	51.8
South Dakota	93.2	90.5	73.2	61.8
Tennessee	90.4	87.6	70.6	55.9
Texas	89.9	84.8	66.6	50.7
Utah	92.0	88.1	71.8	58.2
Vermont	89.7	89.9	72.2	58.2
Virginia	89.2	87.4	67.5	52.6
Washington	90.1	87.0	66.4	52.0
West Virginia	87.1	88.3	66.8	51.4
Wisconsin	90.0	89.7	70.8	57.2
Wyoming	91.4	89.5	73.8	60.4

NOTE: The alternative resolution rate and overall response rate assume that all non-contact cases—telephone numbers for which all call outcomes were “ring, no answer” or busy signals—are not households. The overall response rate is the product of the resolution rate, the age-screener completion rate, and the interview completion rate.

Table I. Covariates used to create nonresponse adjustment cells at different nonresponse adjustment stages

Census region, telephone number listing status, and nonresponse adjustment stage	Covariates ¹
Northeast	
Listed number within area:	
Nonresolution adjustment	Age ² , Education ³ , Minority ⁴
Age-eligibility screener nonresponse adjustment	Owner ⁵ , Minority, Income ⁶ , Age
Unlisted number within area:	
Nonresolution adjustment	Rent ⁷ , MSA ⁸ , Graduate ⁹ , Owner
Age-eligibility screener nonresponse adjustment	Age, MSA, Education
Midwest	
Listed number within area:	
Nonresolution adjustment	Age, Owner, Minority
Age-eligibility screener nonresponse adjustment	Minority, Owner, Income
Unlisted number within area:	
Nonresolution adjustment	MSA, Owner, Minority, Rent
Age-eligibility screener nonresponse adjustment	Age, MSA, Graduate
South	
Listed numbers within area:	
Nonresolution adjustment	Age, Owner, Minority
Age-eligibility screener nonresponse adjustment	Owner, Income, Age, Rent
Unlisted numbers within area:	
Nonresolution adjustment	MSA, Rent, Minority
Age-eligibility screener nonresponse adjustment	Owner, Age, Education
West	
Listed numbers within area:	
Nonresolution adjustment	Age, Owner, Education
Age-eligibility screener nonresponse adjustment	Age, Owner, Education
Unlisted numbers within area:	
Nonresolution adjustment	Minority, Rent, MSA, Graduate
Age-eligibility screener nonresponse adjustment	Age, MSA, Owner, Minority

¹Covariates are listed in order of importance within a group, based on the strength of the relationship between the covariate and observed nonresponse rates.

²Median age in years of the population in the telephone exchange. Median = 36.18.

³Median years of education achieved for the population in the telephone exchange. Median = 13.10.

⁴Percent of population in the telephone exchange who are non-white. Median = 19.70.

⁵Percent of owner-occupied homes in the telephone exchange. Median = 71.40.

⁶Median household income for the telephone exchange, in thousands of dollars. Median = 48.49.

⁷Median monthly rent in the telephone exchange, in dollars. Median = 541.00.

⁸Telephone number is or is not in a Metropolitan Statistical Area (MSA).

⁹Percent of population in the telephone exchange who are college graduates. Median = 23.30.

Table II. Summary statistics for final child-level weights nationally and by state

State	Unweighted sample size	Minimum weight	Maximum weight	Mean weight	Median weight	Sum of weights
National	91,642	0.5	34,724.4	804.9	324.7	73,758,616
Alabama	1,761	9.3	4,574.3	635.6	512.7	1,119,323
Alaska	1,739	2.9	728.2	104.8	84.8	182,287
Arizona	1,769	20.3	6,914.0	937.0	656.4	1,657,543
Arkansas	1,765	52.4	1,964.2	395.8	312.9	698,558
California	1,751	33.9	34,724.4	5,363.8	1,376.7	9,392,086
Colorado	1,801	5.9	4,854.3	659.4	434.3	1,187,560
Connecticut	1,889	36.4	2,206.9	430.7	381.9	813,675
Delaware	1,805	14.8	544.1	111.6	86.8	201,362
District of Columbia	1,801	5.8	311.5	63.2	40.8	113,827
Florida	1,797	13.0	16,266.8	2,235.9	448.8	4,017,889
Georgia	1,782	11.9	9,572.4	1,417.2	915.5	2,525,483
Hawaii	1,822	2.3	848.5	153.6	114.1	279,867
Idaho	1,768	8.0	1,266.0	232.9	189.6	411,741
Illinois	1,932	84.0	10,089.0	1,655.3	1,211.8	3,198,016
Indiana	1,764	24.2	7,329.8	901.7	601.1	1,590,598
Iowa	1,757	4.6	2,914.0	404.1	273.3	710,075
Kansas	1,803	10.7	2,355.1	387.7	322.3	699,044
Kentucky	1,803	58.9	2,753.4	562.1	446.8	1,013,459
Louisiana	1,868	15.3	4,089.5	578.3	393.6	1,080,350
Maine	1,752	6.0	852.1	162.2	134.3	284,110
Maryland	1,773	6.4	4,738.4	768.2	625.0	1,361,936
Massachusetts	1,786	12.9	5,100.8	801.5	640.2	1,431,554
Michigan	1,861	16.1	10,151.7	1,312.6	980.5	2,442,796
Minnesota	1,767	24.0	5,411.8	711.4	490.3	1,257,082
Mississippi	1,918	3.3	2,056.0	394.8	300.7	757,184
Missouri	1,847	15.5	4,615.4	764.5	615.1	1,412,037
Montana	1,788	18.9	664.8	127.5	110.6	227,966
Nebraska	1,827	7.2	1,750.1	250.6	190.3	457,857

Nevada	1,821	0.5	2,906.7	364.8	240.3	664,311
New Hampshire	1,773	2.6	857.2	168.3	145.9	298,439
New Jersey	1,827	12.1	6,685.0	1,121.6	880.6	2,049,175
New Mexico	1,826	5.2	2,077.9	270.3	182.4	493,495
New York	1,793	59.0	14,692.2	2,465.7	1,637.7	4,420,982
North Carolina	1,782	21.7	9,386.6	1,235.6	1,001.9	2,201,857
North Dakota	1,800	4.8	327.5	79.3	71.2	142,697
Ohio	1,765	45.9	11,962.5	1,550.0	1,142.0	2,735,724
Oklahoma	1,816	9.2	2,702.1	497.5	408.6	903,460
Oregon	1,794	15.9	3,589.1	479.0	355.9	859,256
Pennsylvania	1,767	35.7	12,406.5	1,581.3	793.5	2,794,078
Rhode Island	1,756	3.7	899.7	134.3	112.7	235,868
South Carolina	1,885	6.9	3,067.8	561.8	402.7	1,058,919
South Dakota	1,740	7.6	629.3	111.5	95.6	194,049
Tennessee	1,840	65.4	4,084.7	793.3	578.2	1,459,756
Texas	1,805	126.8	19,464.8	3,645.2	1,501.4	6,579,611
Utah	1,747	5.8	3,978.8	465.5	373.8	813,287
Vermont	1,725	2.6	584.2	75.8	51.6	130,729
Virginia	1,774	17.1	5,429.6	1,031.1	862.0	1,829,149
Washington	1,736	11.1	5,636.5	884.6	504.7	1,535,630
West Virginia	1,766	30.5	999.3	221.8	175.3	391,744
Wisconsin	1,860	19.7	3,910.4	706.9	559.5	1,314,848
Wyoming	1,748	2.4	403.9	72.2	62.9	126,287

Table III. Year 2007 guidelines for poverty ranges based on total family members for families in the 48 contiguous states and the District of Columbia

Family Size	Percent of federal poverty level							
	50	100	133	150	185	200	300	400
2	\$6,845	13,690	18,208	20,535	25,327	27,380	41,070	54,760
3	\$8,585	17,170	22,836	25,755	31,765	34,340	51,510	68,680
4	\$10,325	20,650	27,465	30,975	38,203	41,300	61,950	82,600
5	\$12,065	24,130	32,093	36,195	44,641	48,260	72,390	96,520
6	\$13,805	27,610	36,721	41,415	51,079	55,220	82,830	110,440
7	\$15,545	31,090	41,350	46,635	57,517	62,180	93,270	124,360
8	\$17,285	34,570	45,978	51,855	63,955	69,140	103,710	138,280
9	\$19,025	38,050	50,607	57,075	70,393	76,100	114,150	152,200
10	\$20,765	41,530	55,235	62,295	76,831	83,060	124,590	166,120
11	\$22,505	45,010	59,863	67,515	83,269	90,020	135,030	180,040
12	\$24,245	48,490	64,492	72,735	89,707	96,980	145,470	193,960
13	\$25,985	51,970	69,120	77,955	96,145	103,940	155,910	207,880
14	\$27,725	55,450	73,749	83,175	102,583	110,900	166,350	221,800
15	\$29,465	58,930	78,377	88,395	109,021	117,860	176,790	235,720
16	\$31,205	62,410	83,005	93,615	115,459	124,820	187,230	249,640
17	\$32,945	65,890	87,634	98,835	121,897	131,780	197,670	263,560
18	\$34,685	69,370	92,262	104,055	128,335	138,740	208,110	277,480

Table IV. Year 2007 guidelines for poverty ranges based on total family members for families in Hawaii

Family Size	Percent of federal poverty level							
	50	100	133	150	185	200	300	400
2	\$7,875	15,750	20,948	23,625	29,138	31,500	47,250	63,000
3	\$9,875	19,750	26,268	29,625	36,538	39,500	59,250	79,000
4	\$11,875	23,750	31,588	35,625	43,938	47,500	71,250	95,000
5	\$13,875	27,750	36,908	41,625	51,338	55,500	83,250	111,000
6	\$15,875	31,750	42,228	47,625	58,738	63,500	95,250	127,000
7	\$17,875	35,750	47,548	53,625	66,138	71,500	107,250	143,000
8	\$19,875	39,750	52,868	59,625	73,538	79,500	119,250	159,000
9	\$21,875	43,750	58,188	65,625	80,938	87,500	131,250	175,000
10	\$23,875	47,750	63,508	71,625	88,338	95,500	143,250	191,000
11	\$25,875	51,750	68,828	77,625	95,738	103,500	155,250	207,000
12	\$27,875	55,750	74,148	83,625	103,138	111,500	167,250	223,000
13	\$29,875	59,750	79,468	89,625	110,538	119,500	179,250	239,000
14	\$31,875	63,750	84,788	95,625	117,938	127,500	191,250	255,000
15	\$33,875	67,750	90,108	101,625	125,338	135,500	203,250	271,000
16	\$35,875	71,750	95,428	107,625	132,738	143,500	215,250	287,000
17	\$37,875	75,750	100,748	113,625	140,138	151,500	227,250	303,000
18	\$39,875	79,750	106,068	119,625	147,538	159,500	239,250	319,000

Table V. Year 2007 guidelines for poverty ranges based on total family members for families in Alaska

Family Size	Percent of federal poverty level							
	50	100	133	150	185	200	300	400
2	\$8,560	17,120	22,770	25,680	31,672	34,240	51,360	68,480
3	\$10,735	21,470	28,555	32,205	39,720	42,940	64,410	85,880
4	\$12,910	25,820	34,341	38,730	47,767	51,640	77,460	103,280
5	\$15,085	30,170	40,126	45,255	55,815	60,340	90,510	120,680
6	\$17,260	34,520	45,912	51,780	63,862	69,040	103,560	138,080
7	\$19,435	38,870	51,697	58,305	71,910	77,740	116,610	155,480
8	\$21,610	43,220	57,483	64,830	79,957	86,440	129,660	172,880
9	\$23,785	47,570	63,268	71,355	88,005	95,140	142,710	190,280
10	\$25,960	51,920	69,054	77,880	96,052	103,840	155,760	207,680
11	\$28,135	56,270	74,839	84,405	104,100	112,540	168,810	225,080
12	\$30,310	60,620	80,625	90,930	112,147	121,240	181,860	242,480
13	\$32,485	64,970	86,410	97,455	120,195	129,940	194,910	259,880
14	\$34,660	69,320	92,196	103,980	128,242	138,640	207,960	277,280
15	\$36,835	73,670	97,981	110,505	136,290	147,340	221,010	294,680
16	\$39,010	78,020	103,767	117,030	144,337	156,040	234,060	312,080
17	\$41,185	82,370	109,552	123,555	152,385	164,740	247,110	329,480
18	\$43,360	86,720	115,338	130,080	160,432	173,440	260,160	346,880

Table VI. Year 2007 reference value table for additional income cascade questions for households in the 48 contiguous states and District of Columbia

Household size	Reported range of total household income														
	Less than \$7,500	\$7,500- \$9,999	\$10,000- \$12,499	\$12,500- \$14,999	\$15,000- \$17,499	\$17,500- \$19,999	\$20,000- 24,999	\$25,000- \$29,999	\$30,000- \$34,999	\$35,000- \$39,999	\$40,000- \$44,999	\$45,000- \$49,999	\$50,000- \$59,999	\$60,000- \$74,999	\$75,000 and over
2	\$6,800	A	A	\$13,700	B	\$18,200	D	\$27,400	F	F	\$41,100	G	\$54,800	H	H
3	AA	\$8,600	A	A	A	B	\$22,800	D	\$31,800	F	F	F	\$51,500	\$68,700	H
4	AA	AA	A	A	A	A	B	\$27,500	\$31,000	\$38,200	\$41,300	F	F	\$62,000	\$85,000
5	AA	AA	AA	A	A	A	A	B	\$32,100	\$36,200	D	\$48,300	F	\$72,400	\$100,000
6													51,100		
7	AA	AA	AA	\$13,800	A	A	A	\$27,600	B	\$36,700	\$41,400	D	55,200	F	\$85,000/ \$110,000
8	AA	AA	AA	AA	A	A	A	A	\$31,100	B	\$41,400	\$46,600	\$57,500	\$62,200	\$95,000/ \$125,000
9														\$64,000	\$105,000/ \$140,000
10	AA	AA	AA	AA	AA	A	A	A	A	B	B	\$46,000	\$51,900	\$69,100	\$75,000/ \$115,000
11	AA	AA	AA	AA	AA	\$19,000	A	A	A	\$38,100	B	B	\$57,100	\$70,400	\$85,000/ \$125,000
12	AA	AA	AA	AA	AA	AA	A	A	A	A	\$41,500	B	\$55,200	\$62,300	\$90,000/ \$135,000
13	AA	AA	AA	AA	AA	AA	\$22,500	A	A	A	A	B	B	\$67,500	\$64,500/ \$72,700
14	AA	AA	AA	AA	AA	AA	AA	A	A	A	A	\$48,500	B	\$72,700	\$95,000/ \$145,000
15	AA	AA	AA	AA	AA	AA	AA	\$26,000	A	A	A	A	\$52,000	\$69,100	\$105,000/ \$155,000
16	AA	AA	AA	AA	AA	AA	AA	\$27,700	A	A	A	A	\$55,500	\$73,700	\$110,000/ \$165,000
17	AA	AA	AA	AA	AA	AA	AA	AA	A	A	A	A	\$59,000	B	\$120,000/ \$175,000
18	AA	AA	AA	AA	AA	AA	AA	AA	\$31,200	A	A	A	A	\$62,400	\$125,000/ \$185,000

NOTE: When the reported range of household income was included within two or more poverty ranges, additional questions (K11Q59 and K11Q59A) were asked to determine the poverty range for the household. Values within the body of this table represent the border between two poverty ranges. Additional income questions were asked with this value ("Would you say this income was more or less than [value]?") to identify the proper poverty range for the household. Values were rounded to the nearest \$100 if income was below \$75,000 and to the nearest \$5,000 if income was over \$75,000. When income was less than \$20,000, the additional income questions were not asked if the value (i.e., the range border) was less than \$900 from either endpoint of the reported range of household income. Letters rather than values signify that the reported range of household income was entirely within one poverty range. The poverty range for each letter shown is listed in the first bulleted section under "Procedures for Assigning Household Poverty Status" of Appendix VII.

Table VII. Year 2007 reference value table for additional income cascade questions for households in Alaska

Reported Range of Household Income

Household size	Reported Range of Household Income														
	Less than \$7,500	\$7,500-\$9,999	\$10,000-\$12,499	\$12,500-\$14,999	\$15,000-\$17,499	\$17,500-\$19,999	\$20,000-\$24,999	\$25,000-\$29,999	\$30,000-\$34,999	\$35,000-\$39,999	\$40,000-\$44,999	\$45,000-\$49,999	\$50,000-\$59,999	\$60,000-\$74,999	\$75,000 and over
2	AA	\$8,600	A	A	A	B	\$22,800	D	\$31,700	F	F	F	\$51,400	\$68,500	H
3	AA	AA	\$10,700	A	A	A	\$21,500	\$28,600	\$32,200	D	\$42,900	F	F	\$64,400	\$85,000
4	AA	AA	AA	A	A	A	A	B	B	\$38,700	D	\$47,800	\$51,600	F	\$105,000
5	AA	AA	AA	AA	A	A	A	A	B	B	C	D	\$55,800	F	\$120,000
6	AA	AA	AA	AA	AA	A	A	A	A	B	B	C	\$51,800	\$63,900/ \$69,000	\$105,000/ \$140,000
7	AA	AA	AA	AA	AA	\$19,400	A	A	A	\$38,900	B	B	\$51,700/ \$58,300	\$71,900	\$115,000/ \$155,000
8	AA	AA	AA	AA	AA	AA	\$21,600	A	A	A	\$43,200	B	\$57,500	\$64,800	\$85,000/ \$130,000
9	AA	AA	AA	AA	AA	AA	\$23,800	A	A	A	A	\$47,600	B	\$63,300	\$95,000/ \$140,000
10	AA	AA	AA	AA	AA	AA	AA	\$26,000	A	A	A	A	\$51,900	\$69,000	\$105,000/ \$155,000
11	AA	AA	AA	AA	AA	AA	AA	\$28,100	A	A	A	A	\$56,300	\$74,800	\$115,000/ \$170,000
12	AA	AA	AA	AA	AA	AA	AA	AA	A	A	A	A	A	B	\$120,000/ \$180,000
13	AA	AA	AA	AA	AA	AA	AA	AA	\$32,500	A	A	A	A	A	\$125,000/ \$195,000
14	AA	AA	AA	AA	AA	AA	AA	AA	AA	A	A	A	A	\$69,300	\$140,000/ \$210,000
15	AA	AA	AA	AA	AA	AA	AA	AA	AA	\$36,800	A	A	A	\$73,700	\$145,000/ \$220,000
16	AA	AA	AA	AA	AA	AA	AA	AA	AA	\$39,000	A	A	A	A	\$155,000/ \$235,000
17	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	\$41,200	A	A	A	\$165,000/ \$245,000
18	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	\$43,400	A	A	A	\$175,000/ \$260,000

NOTE: When the reported range of household income was included within two or more poverty ranges, additional questions (K11Q59 and K11Q59A) were asked to determine the poverty range for the household. Values within the body of this table represent the border between two poverty ranges. Additional income questions were asked with this value ("Would you say this income was more or less than [value]?") to identify the proper poverty range for the household. Values were rounded to the nearest \$100 if income was below \$75,000 and to the nearest \$5,000 if income was over \$75,000. When income was less than \$20,000, the additional income questions were not asked if the value (i.e., the range border) was less than \$900 from either endpoint of the reported range of household income. Letters rather than values signify that the reported range of household income was entirely within one poverty range. The poverty range for each letter shown is listed in the first bulleted section under "Procedures for Assigning Household Poverty Status" of Appendix VII.

Table VIII. Year 2007 reference value table for additional income cascade questions for households in Hawaii

Reported range of total household income

Household size	Less than \$7,500	\$7,500-\$9,999	\$10,000-\$12,499	\$12,500-\$14,999	\$15,000-\$17,499	\$17,500-\$19,999	\$20,000-\$24,999	\$25,000-\$29,999	\$30,000-\$34,999	\$35,000-\$39,999	\$40,000-\$44,999	\$45,000-\$49,999	\$50,000-\$59,999	\$60,000-\$74,999	\$75,000 and over
2	AA	A	A	A	\$15,800	B	\$23,600	D	\$31,500	F	F	\$47,300	G	\$63,000	H
3	AA	\$9,900	A	A	A	A	B	\$26,300	D	\$36,500	F	F	F	G	\$79,000
4	AA	AA	\$11,900	A	A	A	\$23,800	B	\$31,600	D	\$43,900	\$47,500	F	\$71,300	\$95,000
5	AA	AA	AA	\$13,900	A	A	A	\$27,800	B	\$36,900	\$41,600	D	\$51,300/ \$55,500	F	\$85,000/ \$110,000
6	AA	AA	AA	AA	\$15,900	A	A	A	\$31,800	B	\$42,200	\$47,600	\$58,700	\$63,500	\$95,000/ \$125,000
7	AA	AA	AA	AA	AA	A	A	A	A	B	B	\$47,500	\$53,600	66100/ 71500	\$110,000/ \$145,000
8	AA	AA	AA	AA	AA	AA	A	A	A	A	B	B	\$52,900	\$73,500	\$80,000/ \$120,000
9	AA	AA	AA	AA	AA	AA	\$21,900	A	A	A	\$43,800	B	\$58,200	\$65,600	\$90,000/ \$130,000
10	AA	AA	AA	AA	AA	AA	\$23,900	A	A	A	A	\$47,800	B	63500/ 71600	\$95,000/ \$145,000
11	AA	AA	AA	AA	AA	AA	AA	A	A	A	A	A	\$51,800	\$68,800	\$105,000/ \$155,000
12	AA	AA	AA	AA	AA	AA	AA	\$27,900	A	A	A	A	\$55,800	B	\$110,000/ \$170,000
13	AA	AA	AA	AA	AA	AA	AA	AA	A	A	A	A	A	B	\$120,000/ \$180,000
14	AA	AA	AA	AA	AA	AA	AA	AA	\$31,900	A	A	A	A	\$63,800	\$130,000/ \$190,000
15	AA	AA	AA	AA	AA	AA	AA	AA	\$33,900	A	A	A	A	\$67,800	\$135,000/ \$205,000
16	AA	AA	AA	AA	AA	AA	AA	AA	AA	A	A	A	A	\$71,800	\$145,000/ \$215,000
17	AA	AA	AA	AA	AA	AA	AA	AA	AA	\$37,900	A	A	A	A	\$150,000/ \$230,000
18	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	A	A	A	A	\$160,000/ \$240,000

NOTE: When the reported range of household income was included within two or more poverty ranges, additional questions (K11Q59 and K11Q59A) were asked to determine the poverty range for the household. Values within the body of this table represent the border between two poverty ranges. Additional income questions were asked with this value ("Would you say this income was more or less than [value]?") to identify the proper poverty range for the household. Values were rounded to the nearest \$100 if income was below \$75,000 and to the nearest \$5,000 if income was over \$75,000. When income was less than \$20,000, the additional income questions were not asked if the value (i.e., the range border) was less than \$900 from either endpoint of the reported range of household income. Letters rather than values signify that the reported range of household income was entirely within one poverty range. The poverty range for each letter shown is listed in the first bulleted section under "Procedures for Assigning Household Poverty Status" of Appendix VII.

Table IX. State-specific insurance program names used for questions about Medicaid and the State Children's Health Insurance Program

State	Program Name
Alabama	Patient 1st Program, or ALL Kids
Alaska	Denali KidCare
Arizona	AHCCCS, or KidsCare
Arkansas	ConnectCare, or ARKids First
California	Medi-Cal, or Healthy Families Program
Colorado	MAC Card or the Primary Care Physician Program, or Child Health Plan Plus
Connecticut	the HUSKY Plan
Delaware	Diamond State Health Plan, or Delaware Healthy Children Program
District of Columbia	DC Healthy Families
Florida	Florida KidCare, which includes the Healthy Kids and Medi-Kids programs
Georgia	Georgia Better Health Care Program and Georgia Healthy Families, or PeachCare for Kids
Hawaii	Hawaii-QUEST
Idaho	Healthy Connections, or Children's Health Insurance Program, or CHIP
Illinois	All Kids
Indiana	the Hoosier Healthwise program
Iowa	MediPASS, or HAWK-I (Healthy and Well Kids in Iowa)
Kansas	HealthConnect Kansas, or HealthWave
Kentucky	KyHealth Choices or KENPAC, or Kentucky Children's Health Insurance Program (K-CHIP)
Louisiana	CommunityCARE program or the Louisiana Children's Health Insurance Program
Maine	Maine Care
Maryland	Medical Assistance Program, or HealthChoice, or Maryland Children's Health Program
Massachusetts	MassHealth
Michigan	Healthy Kids Program, or MI-Child Program
Minnesota	Medical Assistance or MinnesotaCare
Mississippi	Mississippi Children's Health Insurance Program
Missouri	MC-Plus and MC-Plus For Kids
Montana	Passport to Health program, or Montana Child Health Insurance Plan, or CHIP
Nebraska	Kids Connection or the Nebraska Health Connection program
Nevada	Nevada Check Up
New Hampshire	Healthy Kids Gold or Health Kids Silver
New Jersey	New Jersey FamilyCare
New Mexico	SALUD!, or New MexiKids
New York	Child Health Plus
North Carolina	Carolina ACCESS, or North Carolina Health Choice for Children
North Dakota	Healthy Steps program
Ohio	Healthy Start and Healthy Families
Oklahoma	SoonerCare
Oregon	Oregon Health Plan, or Oregon Children's Health Insurance Program
Pennsylvania	HealthChoices, or the ACCESS Card, or CHIP, the Children's Health Insurance Program
Rhode Island	Rite Care
South Carolina	Partners for Healthy Children program
South Dakota	the PRIME program, or the Child Health Insurance Program
Tennessee	TennCare
Texas	State of Texas Access Reform program, or STAR program, or Texas Children's Health Insurance Program
Utah	Utah Children's Health Insurance Program, or CHIP
Vermont	Dr. Dynasaur
Virginia	Medallion program, or FAMIS (Family Access to Medical Insurance Security Plan)
Washington	Healthy Options program or Basic Health Plus, or Washington State's Children's Health Insurance Program
West Virginia	West Virginia Physician Assured Access System, or the Mountain Health Trust program, or West Virginia Children's Health Insurance Program
Wisconsin	BadgerCare
Wyoming	EqualityCare, or Wyoming KidCare

Table X. Frequencies of disposition codes

Disposition code by name	Disposition category	Frequency	Percent of total
Total number of telephone lines in sample		2,806,416	100.00
Not attempted	UH	7	0.00
No contact	UH	191,956	6.84
Answering machine—residential status unknown	UH	82,561	2.94
Spanish case—residential status unknown	UH	210	0.00
Appointment at introduction—residential status unknown	UH	1,628	0.06
Callback at introduction—residential status unknown	UH	10,475	0.37
Hang-up during introduction—residential status unknown	UH	87,920	3.13
Refusal at introduction—residential status unknown	UH	71,133	2.53
Other Introduction—residential status unknown	UH	82	0.00
Answering machine—known household	UO	2,317	0.08
Appointment—known household (NIS screening pending)	UO	8,468	0.30
Callback—known household (NIS screening pending)	UO	23,435	0.84
Refusal—known household (NIS screening pending)	UO	25,110	0.89
Other—known household (NIS screening pending)	UO	488	0.00
Appointment prior to NSCH Item S_UNDR18 (NSCH screening pending)	UO	621	0.00
Callback prior to NSCH Item S_UNDR18 (NSCH screening pending)	UO	3,761	0.13
Refusal prior to NSCH Item S_UNDR18 (NSCH screening pending)	UO	9,530	0.34
Other prior to NSCH Item S_UNDR18 (NSCH screening pending)	UO	312	0.00
Prefinalized Do Not Call List	UO	9	0.00
Minor HH	X	2,843	0.10
Age Ineligible	X	377,287	13.44
NIS-level appointment (NIS eligible)	R	114	0.00
NIS-level callback (NIS eligible)	R	52	0.00

NIS-level refusal (NIS eligible)	R	1,159	0.00
NIS-level other (NIS eligible)	R	9	0.00
NIS-finalized eligible for SLAITS redialing, recontact unsuccessful	R	1,748	0.06
Teen-level appointment (Teen screening pending)	R	41	0.00
Teen-level callback (Teen screening pending)	R	102	0.00
Teen-level refusal (Teen screening pending)	R	317	0.00
Teen-level other (Teen screening pending)	R	10	0.00
Teen-level appointment (Teen eligible)	R	6	0.00
Teen-level callback (Teen eligible)	R	5	0.00
Teen-level refusal (Teen eligible)	R	121	0.00
Teen-level other (Teen eligible)	R	1	0.00
Teen-finalized eligible for SLAITS redialing, recontact unsuccessful	R	237	0.00
Appointment prior to NSCH Item K8Q12	R	7,602	0.27
Callback prior to NSCH Item K8Q12	R	9,673	0.34
Refusal prior to NSCH Item K8Q12	R	21,762	0.78
NIS-finalized eligible for SLAITS redialing, ended prior to NSCH Item K8Q12	R	274	0.00
Teen-finalized eligible for SLAITS redialing, ended prior to NSCH Item K8Q12	R	38	0.00
Other prior to NSCH Item K8Q12	R	463	0.00
Appointment-partial interview	P	309	0.00
Callback-partial interview	P	350	0.00
Refusal-partial interview	P	413	0.00
Other-partial interview	P	5	0.00
NIS-finalized eligible for SLAITS redialing-partial interview	P	7	0.00
Teen-finalized eligible for SLAITS redialing-partial interview	P	1	0.00
Completed household interview	I	71,388	2.54
Converted household interview from refusal	I	18,712	0.67
Converted household interview, NIS-finalized eligible for SLAITS redialing	I	407	0.00
Converted household interview, Teen-finalized eligible for SLAITS redialing	I	50	0.00
3 or more fax/modem prior to contact	Z	35,575	1.27
3 or more fast busy prior to contact	Z	14,810	0.53

3 or more other technological problem prior to contact	Z	630	0.00
2 or more not in service	Z	250,155	8.91
Nonworking	Z	22,451	0.80
Number changed	Z	806	0.00
Not residential	Z	119,746	4.27
GENESYS-resolved telephone numbers (nonworking, business, and fax or modem)	Z	1,326,714	47.27
NOTE: 0.00 Quantity more than zero, but less than 0.05			

Table XI. Unweighted response rate calculations

Disposition categories and response rates	Frequency or calculated rate	Code or formula
Summary of disposition categories		
Not resolved as residential or non- residential	445,972	UH
Known household, age eligibility undetermined	74,051	UO
Out of scope (business, nonworking, fax, or modem)	1,770,887	Z
Age-screened household, no child in range	380,130	X
Known age-eligible household	43,734	R
Eligible household, partially completed interview	1,085	P
Eligible household, completed interview	90,557	I
Total	2,806,416	...
Calculation of response rates		
Resolution rate (RR)	84.1	$(I+P+R+X+UO+Z) / (I+P+R+X+UO+Z+UH)$
Screenener completion rate (SCR)	87.4	$(I+P+R+X) / (I+P+R+X+UO)$
Interview completion rate (ICR)	67.7	$(I+P) / (I+P+R)$
Overall response rate	49.8	$(ICR)(SCR)(RR)$
...Not applicable		

Table XII. Completion rates by incentive eligibility type

	Initial incentive model		Expanded incentive effort, \$15			
	Tier 1, \$10	Tier 2, \$15	Overall	No refusal	One refusal	Total, any incentive
Eligible cases	59,952	30,121	9,177	4,589	4,588	66,365
Completed cases	19,087	9,408	3,293	1,881	1,412	31,788
Completion rate	31.8	31.2	35.9	41.0	30.8	47.9

Table XIII. Completion rates by incentive eligibility type and quarter

Incentive eligibility type	Quarter 1 2007	Quarter 2 2007	Quarter 3 2007	Quarter 4 2007	Quarter 1 2008
	banked sample				
Initial incentive model					
Tier 1, \$10	34.0	28.9	32.2	32.4	33.1
Tier 2, \$15	32.7	29.7	30.6	32.1	32.7
Expanded incentive effort, \$15					
Overall	41.5	26.4	32.7	39.4	42.0
No refusal	47.5	28.0	38.0	46.3	47.4
One refusal	35.0	24.9	27.7	32.2	36.5

Table XIV. Parameter estimates for a linear regression model predicting reported household income values and relative odds from a logistic regression model predicting whether income was not reported

Variable	Parameters of linear regression model to predict household income values				Relative odds of having missing income based on a logistic regression model	
	Estimate	Standard error	t- statistic	p-value	Odds ratio	95% confidence interval
Intercept	-6.71	0.67	-10.06	<0.01
Child is black	-0.50	0.04	-11.88	<0.01	1.02	0.94-1.11
Child is white	0.05	0.03	1.69	0.09	1.13	1.06-1.21
Education of respondent	0.81	0.01	62.81	<0.01	0.94	0.91-0.96
Household size (base-10 log transformed)	3.39	0.15	22.87	<0.01	1.45	1.09-1.93
MSA status	-0.54	0.02	-23.08	<0.01	1.20	1.14-1.26
S_UNDR18	-0.55	0.04	-15.64	<0.01	0.88	0.83-0.95
K1Q01	-0.02	0.02	-1.20	0.23	0.98	0.94-1.02
K1Q03	-0.47	0.05	-9.01	<0.01	1.33	1.21-1.46
K2Q01	-0.19	0.02	-12.69	<0.01	1.04	1.01-1.07
K2Q13	0.06	0.04	1.57	0.12	0.86	0.79-0.92
K2Q19	-0.18	0.05	-3.77	<0.01	1.08	0.98-1.18
K2Q22	0.03	0.05	0.68	0.50	0.90	0.81-1.00
K2Q32A	-0.03	0.06	-0.46	0.64	0.93	0.82-1.06
K2Q33A	0.07	0.05	1.41	0.16	0.97	0.87-1.08
K2Q34A	0.00	0.06	-0.01	0.99	1.03	0.91-1.16
K2Q37A	0.07	0.05	1.46	0.14	0.94	0.86-1.04
K2Q40A	0.07	0.03	2.27	0.02	0.97	0.92-1.03
K2Q41A	0.22	0.14	1.54	0.12	0.98	0.73-1.30
K2Q42A	-0.07	0.10	-0.70	0.48	1.04	0.86-1.26
K2Q43A	-0.05	0.06	-0.85	0.39	1.04	0.92-1.17
K2Q44A	0.02	0.08	0.19	0.85	1.12	0.97-1.31
K2Q45A	0.09	0.06	1.62	0.11	1.05	0.94-1.17
K2Q46A	0.06	0.07	0.87	0.38	1.12	0.99-1.27

K3Q01	0.70	0.04	17.15	<0.01	0.85	0.79-0.92
K3Q21A	0.61	0.02	24.75	<0.01	0.94	0.90-0.99
K4Q20	-0.03	0.01	-4.70	<0.01	1.00	0.99-1.01
K4Q21	0.14	0.01	18.34	<0.01	1.03	1.01-1.04
K4Q27	-0.41	0.04	-9.41	<0.01	0.84	0.77-0.92
K5Q10	0.07	0.03	2.50	0.01	0.96	0.91-1.01
K9Q00	-0.43	0.04	-11.36	<0.01	1.03	0.96-1.10
K9Q40	-0.19	0.02	-7.87	<0.01	0.72	0.68-0.75
K10Q12	-0.04	0.03	-1.49	0.14	0.94	0.90-0.99
K10Q13	0.22	0.02	9.99	<0.01	0.98	0.93-1.02
K10Q14	-0.01	0.03	-0.17	0.86	0.94	0.89-1.00
K10Q23	-0.18	0.03	-5.36	<0.01	0.84	0.78-0.90
K10Q40	0.21	0.01	15.20	<0.01	0.97	0.95-1.00
K11Q01	-0.20	0.04	-5.55	<0.01	0.90	0.84-0.97
K11Q33	0.26	0.07	3.86	<0.01	0.83	0.74-0.93
K11Q38	1.05	0.13	7.96	<0.01	1.19	0.94-1.49
K11Q50	1.17	0.04	29.81	<0.01	0.66	0.62-0.71
K11Q60	-0.60	0.06	-10.04	<0.01	1.24	1.12-1.38
K11Q61	-1.70	0.04	-39.43	<0.01	1.31	1.21-1.42
K11Q62	-1.92	0.03	-60.00	<0.01	0.98	0.92-1.04
K11Q70	0.77	0.03	24.31	<0.01	1.19	1.12-1.26
K11Q76A (days)	-0.12	0.02	-7.62	<0.01	0.95	0.92-0.98
Mean household income by state	1.65	0.07	22.74	<0.01
Standard deviation of income by state	-0.24	0.02	-13.58	<0.01

... Category not applicable.

Table XV. Unweighted and weighted estimates of the frequency and prevalence of children with excellent or very good health

State	Total unweighted number of children	Total weighted estimate of number of children	Unweighted number of children with excellent or very good health	Weighted estimate of number of children with excellent or very good health	Standard error of weighted estimate of number of children with excellent or very good health	Percent of children who have excellent or very good health	Standard error of percent of children who have excellent or very good health
National	91,642	73,758,616	80,456	62,216,654	458,838	84.35	0.347
Alabama	1,761	1,119,323	1,524	947,115	35,054	84.61	1.317
Alaska	1,739	182,287	1,546	162,052	5,624	88.90	1.057
Arizona	1,769	1,657,543	1,505	1,337,038	50,813	80.66	1.549
Arkansas	1,765	698,558	1,507	574,474	18,031	82.24	1.262
California	1,751	9,392,086	1,427	7,300,335	359,746	77.73	1.952
Colorado	1,801	1,187,560	1,597	1,001,551	38,710	84.34	1.580
Connecticut	1,889	813,675	1,706	717,021	20,863	88.12	1.030
Delaware	1,805	201,362	1,555	170,122	5,445	84.49	1.168
District of Columbia	1,801	113,827	1,558	95,602	3,269	83.99	1.212
Florida	1,797	4,017,889	1,528	3,571,983	166,312	88.90	1.319
Georgia	1,782	2,525,483	1,588	2,177,204	83,832	86.21	1.381
Hawaii	1,822	279,867	1,599	241,938	7,915	86.45	1.155
Idaho	1,768	411,741	1,549	354,123	11,380	86.01	1.115
Illinois	1,932	3,198,016	1,650	2,715,176	86,471	84.90	1.131
Indiana	1,764	1,590,598	1,505	1,357,295	49,965	85.33	1.206
Iowa	1,757	710,075	1,577	620,713	21,836	87.42	1.229
Kansas	1,803	699,044	1,580	596,113	19,451	85.28	1.168
Kentucky	1,803	1,013,459	1,567	874,230	27,097	86.26	1.036
Louisiana	1,868	1,080,350	1,591	868,362	31,348	80.38	1.510
Maine	1,752	284,110	1,608	258,153	8,054	90.86	0.934
Maryland	1,773	1,361,936	1,619	1,227,016	41,221	90.09	1.085
Massachusetts	1,786	1,431,554	1,620	1,264,383	42,866	88.32	1.190
Michigan	1,861	2,442,796	1,639	2,110,493	75,077	86.40	1.317
Minnesota	1,767	1,257,082	1,617	1,148,535	41,370	91.37	0.987
Mississippi	1,918	757,184	1,619	622,709	20,157	82.24	1.208
Missouri	1,847	1,412,037	1,633	1,232,395	39,520	87.28	1.059
Montana	1,788	227,966	1,604	201,191	5,953	88.25	1.022
Nebraska	1,827	457,857	1,644	395,150	13,584	86.30	1.317
Nevada	1,821	664,311	1,486	530,170	20,737	79.81	1.484
New Hampshire	1,773	298,439	1,636	270,414	8,373	90.61	0.995
New Jersey	1,827	2,049,175	1,633	1,756,053	58,927	85.70	1.268
New Mexico	1,826	493,495	1,558	420,042	16,220	85.12	1.194
New York	1,793	4,420,982	1,520	3,684,697	128,478	83.35	1.313
North Carolina	1,782	2,201,857	1,554	1,905,359	67,353	86.53	1.214
North Dakota	1,800	142,697	1,640	129,226	3,647	90.56	0.849
Ohio	1,765	2,735,724	1,586	2,363,238	84,518	86.38	1.390
Oklahoma	1,816	903,460	1,586	774,945	25,340	85.78	1.087
Oregon	1,794	859,256	1,562	741,477	25,744	86.29	1.171
Pennsylvania	1,767	2,794,078	1,553	2,478,407	96,904	88.70	1.184
Rhode Island	1,756	235,868	1,564	203,057	7,039	86.09	1.304
South Carolina	1,885	1,058,919	1,642	894,640	29,473	84.49	1.190
South Dakota	1,740	194,049	1,580	174,735	5,456	90.05	0.943
Tennessee	1,840	1,459,756	1,622	1,230,196	40,015	84.27	1.267
Texas	1,805	6,579,611	1,444	5,141,109	229,426	78.14	1.633
Utah	1,747	813,287	1,592	732,234	24,699	90.03	1.058
Vermont	1,725	130,729	1,604	120,723	4,276	92.35	0.938
Virginia	1,774	1,829,149	1,583	1,608,914	52,654	87.96	1.075
Washington	1,736	1,535,630	1,516	1,316,713	51,968	85.74	1.303
West Virginia	1,766	391,744	1,544	339,619	10,404	86.69	1.039
Wisconsin	1,860	1,314,848	1,643	1,147,422	36,436	87.27	1.082
Wyoming	1,748	126,287	1,546	110,793	3,671	87.73	1.020

NOTE: Denominator includes children for whom health status was not reported because the respondent did not know or refused to answer the health status question. Estimates are derived from the publicly released data file.