

Appendix F

**2021 National Survey of Children's Health
Methodology Report**

October 14, 2022

2021 National Survey of Children's Health

Methodology Report

Contents

Abstract.....	5
Objectives.....	5
Methods.....	5
Results.....	5
Introduction	6
Survey History	7
Frame, Sample, and Selected Child Subsample	8
Frame and Sample Selection.....	8
Selected Child Subsample	9
Data Collection.....	11
Survey Content.....	11
2021 Content Changes.....	12
Data Collection Instruments	13
Treatment Groups	17
Mailing Contents and Schedule.....	18
Response Analysis	21
Response Rates	21
Item-Level Response.....	23
Treatment Groups and Response.....	25
Data Processing.....	28
Unduplication.....	28
Paper to Web Standardization	29
Data Edits.....	29
Recoded and Standardized Variables	31
Missing Values and Imputation	35
Suppressed Variables.....	37
Geography Variables.....	38
Weighting Specifications	40
Overview	40
Population Controls.....	43
Limitations.....	44
Estimation, Hypothesis Testing, and Data Use Guidelines	46
Variance Estimation.....	46
Statistical Testing for Overlapping Groups.....	47
Combining Data across Survey Years.....	48
Confidentiality	48
Guidelines for Data Use	48
Supporting Material	50
References.....	50

Attachment A: 2021 Estimated State-Level Sample Sizes	51
Attachment B: Probabilities for Selected Child	53
Attachment C: Weighted Response Rates by State	55

Table of Figures

<i>Table 1. Web Submission Times (in Minutes)</i>	14
<i>Table 2. TQA Purpose Codes used in ATAC System</i>	16
<i>Table 3. Production Mailout Schedule</i>	18
<i>Table 4. Topical Mailings and Topical Mailing Groups</i>	20
<i>Table 5. 2021 Final Dispositions (Unweighted)</i>	21
<i>Table 6. 2021 NSCH Weighted Response Rates</i>	23
<i>Table 7. Lowest Item-Level Response Rates</i>	24
<i>Table 8. Data Collection Costs and Completed Questionnaires by Screener Incentive</i>	25
<i>Table 9. Data Collections Costs and Returns by Topical Incentive</i>	26
<i>Table 10. Response Odds Ratios (Incentive versus No Incentive) by Education, Race, and Income</i>	27
<i>Table 11. Unduplication Criteria for both Web and Paper Returns</i>	28
<i>Table 12. Unduplication Criteria for Two Paper Returns</i>	29
<i>Table 13. Standardized Variables</i>	31
<i>Table 14. Derived and Recoded Variables</i>	31
<i>Table 15. Imputed Variables and Their Imputation Flags</i>	36
<i>Table 16. Suppressed Variables</i>	37
<i>Table 17. List of Geography Variables</i>	38
<i>Table 18. Geographies Identified at the Intersections</i>	39
<i>Table 19. Collapsed Dimensions of Final Raking and Affected States</i>	44
<i>Table A-1: Address Sample Size and Strata Distribution by State</i> ²⁰	51
<i>Table B-1: Household Type Assignment from the Values of Five Screener Variables</i> ²¹	53
<i>Table C-1. Weighted Response Rates by State</i> ²²	55

Abstract

Objectives

This report details the development, plan, and operation of the 2021 National Survey of Children's Health (NSCH). This survey is designed to provide national and state-level estimates on key indicators of the health and well-being of children, their families and communities, as well as information about the prevalence and impact of special health care needs. Funding and direction for this survey was provided by the Health Resources and Services Administration's Maternal and Child Health Bureau (HRSA MCHB) within the U.S. Department of Health and Human Services. The U.S. Census Bureau conducted the survey on behalf of HRSA MCHB.

Methods

The 2021 NSCH used a national sample of 300,000 addresses. During data collection, a screener questionnaire was used to identify households with children and roster children in the household. The screener questionnaire also included a battery of questions to identify children with special health care needs. One child was randomly selected from each eligible household, and that child was the subject of a more detailed topical questionnaire. Responses to the screener and topical questionnaires were collected, processed, and published in the Screener Public Use File and Topical Public Use File.

Results

The weighted Overall Response Rate for the 2021 NSCH was 40.3%. A total of 106,000¹ screener questionnaires were completed, and of those 62,010 were eligible for topical questionnaire follow-up. Of those topical-eligible households, 50,892 completed a topical interview. Weighted estimates from the Topical file generalize to state and national resident child populations. Weighted estimates from the Screener file generalize to state and national resident child populations (using the child weight) and households with children by state and nationally (using the household weight).

¹ Rounded to the nearest thousand in accordance with the U.S. Census Bureau disclosure avoidance practices.

Introduction

The 2021 National Survey of Children’s Health (NSCH) was conducted by the U.S. Census Bureau for the Health Resources and Services Administration, Maternal and Child Health Bureau (HRSA MCHB) within the U.S. Department of Health and Human Services (HHS). As stated in the Office of Management and Budget Clearance Package, the purpose of the NSCH is to “collect information on factors related to the well-being of children, including access to and quality of health care, family interactions, parental health, school and after-school experiences, and neighborhood characteristics.”² This document details the objectives, methodologies, and results of the 2021 NSCH into seven sections.

- *Survey History.* The 2021 NSCH was the fifth annual production cycle following the redesign and merging of the previous NSCH and National Survey of Children with Special Health Care Needs (NS-CSHCN).
- *Frame, Sample and Selected Child Subsample.* A screener questionnaire identified households with children and rostered the children in those households. A topical questionnaire collected detailed information about one child selected at random from the household.
- *Data Collection.* Data were collected using a two-stage paper survey instrument and a single-stage web-based survey instrument. This section discusses treatment groups, mail schedule and data capture methods for the web, paper, and telephone questionnaire assistance operations.
- *Response Analysis.* This section discusses the calculation of response rates along with analyses of survey breakoffs, item nonresponse, and treatment group comparisons.
- *Data Processing.* Web and paper survey responses were cleaned for analysis, including unduplication of responses, edits for data quality, creating standardized and derived variables, and imputation of missing values.
- *Weighting Specifications.* Weights allow for estimates to be generalized to state and national child resident populations (Screener and Topical file) and households with children (Screener file).
- *Estimation, Hypothesis Testing, and Data Use Guidelines.* A discussion of the best practices for data users and limitations of the 2021 NSCH.

² The Office of Management and Budget Clearance Package is available at https://www.reginfo.gov/public/do/PRAViewDocument?ref_nbr=202102-0607-001

Survey History

The Health Resources and Services Administration's Maternal and Child Health Bureau (HRSA MCHB), within the U.S. Department of Health and Human Services (HHS), has sponsored the National Survey of Children's Health (NSCH)³ and its companion survey, the National Survey of Children with Special Health Care Needs (NS-CSHCN),⁴ since 2001. HRSA MCHB has provided funding and direction for the two periodic surveys to provide both national and state estimates of key indicators of child health and well-being for children ages 0-17 years.

Together, these surveys provided critical data on key measures of child health; the presence and impact of special health care needs; health care access, utilization, and quality; and the family and community factors that impact child and adolescent health and well-being. Both surveys were fielded three times (NS-CSHCN 2001, 2005-06, and 2009-10; NSCH 2003, 2007, and 2011-12) as modules of the State and Local Area Integrated Telephone Survey (SLAITS) system by the Centers for Disease Control and Prevention's National Center for Health Statistics. As part of the SLAITS system, the surveys utilized a random-digit-dial sample of landline telephone numbers, with cell-phone supplementation in the last year of administration for both surveys.

In 2015, HRSA MCHB redesigned the NSCH and the NS-CSHCN into a single combined survey that utilized an address-based sampling frame. When this newly consolidated survey was first fielded in 2016 it incorporated questions from the former surveys and retained the NSCH name. The U.S. Census Bureau now conducts the NSCH annually on behalf of HRSA MCHB and HHS under Title 13, United States Code, Section 8(b), which allows the Census Bureau to conduct surveys on behalf of other agencies.

³ Blumberg SJ, Foster EB, Frasier AM, et al. 2012. Design and Operation of the National Survey of Children's Health, 2007. National Center for Health Statistics. *Vital Health Stat*, 1(55).

http://www.cdc.gov/nchs/data/series/sr_01/sr01_055.pdf

⁴ Bramlett MD, Blumberg SJ, Ormson AE, et al. 2014. Design and Operation of the National Survey of Children with Special Health Care Needs, 2009–2010. National Center for Health Statistics. *Vital Health Stat*, 1(57).

http://www.cdc.gov/nchs/data/series/sr_01/sr01_057.pdf

Frame, Sample, and Selected Child Subsample

The 2021 NSCH sampled approximately 300,000 addresses to participate in the survey. One child from each household with children was selected, or subsampled, to be the subject of the topical questionnaire. This section covers the design of the sample and subsample.

Frame and Sample Selection

The 2021 NSCH used an address-based sample selected from an extract of the Census Bureau's Master Address File (MAF)⁵. It covers the 50 states and the District of Columbia⁶. The sample frame uses administrative records-based flags to identify four mutually exclusive strata:

- Stratum 1A and 1B: Addresses that are explicitly linked to children using administrative records are assigned to Stratum 1. Approximately 80% of these addresses are households with children. Within Stratum 1, if a linked child is 5 years old or younger, the address is assigned to Stratum 1A; otherwise, the address is assigned to Stratum 1B.
- Stratum 2a: Addresses that are probabilistically linked to children using administrative records and block group characteristics. Approximately 15% of these addresses are households with children.
- Stratum 2b: The remaining addresses. Less than 5% of these addresses are households with children.

Addresses assigned to Stratum 1 are explicitly linked to a child record either directly or through a parent using administrative and survey records. Data sources include:

- Social Security applications and the Census Numident
- IRS 1040s and 1099s
- Medicare Enrollment Database (MEDB)
- Indian Health Service database (IHS)
- Selective Service System (SSS)
- Public Indian Housing Information Center (PIC) and Tenant Rental Assistance Certification System (TRACS) data from the Department of Housing and Urban Development (HUD)
- National Change of Address data from the US Postal Service
- American Community Survey and CPS-ASEC, and the 2010 Census Unedited files (for parent-child links)

Approximately 38 million unique addresses were linked to at least one child record and assigned to Stratum 1. Stratum 1 was then subdivided into Strata 1a and 1b. Stratum 1a addresses were those most likely to have young children (0 – 5 years old). Stratum 1A was sampled at a higher rate than Stratum 1B in order to increase representation of young children in the NSCH.

⁵ The MAF is a Title 13 data source, and all data collected are confidential under 13 U.S.C. Section 9. All access to Title 13 data from this survey is restricted to Census Bureau employees and those holding Census Bureau Special Sworn Status pursuant to 13 U.S.C. Section 23(c).

⁶ Hereafter, 'state' will include the District of Columbia.

The remaining addresses were then subdivided into Strata 2a and 2b. All Stratum 2 addresses were assigned a probability of child presence using administrative records and small-area geographic characteristics. Beginning with those addresses with the lowest probability of children presence, addresses were assigned to Stratum 2b by state until the stratum represented at most 5% of households with children in that state (as reported in the 2019 American Community Survey). All other addresses were assigned to Stratum 2a, and Strata 1 and 2a combined represented 95% of households with children in each state.

To increase the efficiency of the sample, Stratum 1a addresses were sampled at the highest rate, then Stratum 1b. Stratum 2a addresses were sampled at a lower rate and addresses in stratum 2b were excluded from sampling. For the sample selection:

- The sampling rates by strata in each state were optimized to maximize the number of households without children in each state without compromising the reliability of survey estimates. Nationally, 70% of addresses came from Stratum 1, 40% from Stratum 1A and 30% from Stratum 1B, and 30% from Stratum 2a.
- The addresses within each state were first sorted by strata, then organized into two groups by the block group⁷ poverty rate to ensure states had proportional representation of addresses in high poverty areas selected for the sample.
- The sample was distributed across states to produce a roughly equal number of completed interviews per state. Four states included an oversample (see Attachment A) to increase the number of interviews completed in those states.
- To minimize respondent burden, addresses can be selected only once in any five-year period.

Selected Child Subsample

The screener questionnaire collects information on the presence of children within the household, child demographic information, and basic questions about each child's health.⁸ One child is selected from the completed screener, and one of three age-based topical questionnaires is provided to the household based on the sampled child's age:

- NSCH-T1: children aged 0 through 5,
- NSCH-T2: children aged 6 through 11, or
- NSCH-T3: children aged 12 through 17

The probability of selection for a child is based on the number of children in the household, the special health care needs status, and the age of the child. When appropriate, an 80% oversample is applied to

⁷ A Census block group is a geographical unit with 600 to 3,000 population. Census blocks are grouped into block groups; block groups, in turn, are grouped into Census tracts. The block group is the smallest scale geographical unit for which the Census Bureau publishes sample statistics, i.e., estimates based on a sample of residents in the block group. Consequently, it is the smallest scale geographical unit that could be used for this exercise.

⁸ Bethell CD, Read D, Neff J, Blumberg SJ, Stein RE, Sharp V, Newacheck PW. 2002. "Comparison of the Children with Special Health Care Needs Screener to the Questionnaire for Identifying Children with Chronic Conditions — Revised." *Ambulatory Pediatrics*, Jan-Feb 2(1): 49-57.

children with special health care needs and a 60% oversample to young children (ages 0-5).⁹ See Attachment B for more details.¹⁰

For mailed-in screener responses, the appropriate topical questionnaire is mailed to the household, and mail materials indicate which child has been selected. In the web-based instrument, the child's reported age is used to navigate the respondent to age-appropriate survey items.

⁹ The 80% oversample is applicable only for those households with both CSHCN and Non-CSHCN present. The 60% age-based oversample is applicable when the conditions of the CSHCN oversample are not met and there are both young (ages 0-5) and older (ages 6-17) children present.

¹⁰ Eligible children in a household are sorted first by special health care needs status (CSHCN then Non-CSHCN) and then by age (youngest to oldest). Additionally, children with the same special health care needs status and age are sorted by name. In households with four or more eligible children, children are sorted first on special health care needs status, then alphabetically by name, and then by age.

Data Collection

Data collection efforts for the 2021 National Survey of Children’s Health (NSCH) began on June 25, 2021 and continued until January 14, 2022. The 2021 NSCH retained a two-phase data collection approach: (1) an initial household screener to assess the presence, basic demographic characteristics, and special health care needs status of any children in the home; and (2) a substantive topical questionnaire to be completed by a parent or caregiver of the selected child. The data collection methodology employed strategies to increase response, including clear and concise question wording, providing response mode options, cash incentives and other treatments.

This section covers survey content and 2021 content changes, data collection instruments, and the data collection process.

Survey Content

Questionnaires were designed to encourage cooperation by prospective respondents, enhance respondent comprehension, and make instructions clear and simple. Questions were developed and grouped by subject area to create logical, clear questionnaires with concrete question wording and simple grammar.

The screener questionnaire consisted of two sections. The first section contained four questions about the presence of children in the home, the primary language spoken, and home tenure (rent or own). The next section contained detailed questions about the demographics and health of children in the household.

There were three different topical questionnaires tailored to three child age groups: NSCH-T1 for 0 to 5-year-old children, NSCH-T2 for 6 to 11-year-old children, and NSCH-T3 for 12 to 17-year-old children. All three questionnaires contained 11 sections about the child, their family and neighborhood, but the specific questions were tailored to be relevant to children in that age range. Copies of the screener and topical questionnaires can be found at <https://www.census.gov/programs-surveys/nsch/technical-documentation/questionnaires.html>.

Section A. This Child’s Health

Current or lifelong physical, mental, behavioral, learning, or developmental conditions, and the impact of these conditions on the child’s activities.

Section B: This Child as an Infant

Birth-related questions including birth weight, breastfeeding, and use of formula. Infant feeding questions are only included on NSCH-T1.

Section C: Health Care Services

Health care providers and the child’s need for and use of medical, dental, mental, and specialized health services in the last 12 months.

Section D: Experience with This Child's Health Care Providers

Frequency of care and satisfaction with the child's health care providers, and how the child's doctor or health care providers worked with the child. NSCH-T3 includes questions about the child's preparation for transition into adult health care.

Section E: This Child's Health Insurance Coverage

Status and adequacy of health insurance coverage, including any gaps in health insurance coverage in the past 12 months.

Section F: Providing for this Child's Health

Cost of health care in the past 12 months and time spent providing and arranging for the child's health care.

Section G: This Child's Learning/Schooling and Activities

Early language development and learning for children ages 1 to 5 years. For children ages 6 to 17 years, experiences at school, participation in organized activities, and physical activities.

Section H: About You and This Child

Daily life and household activities, including the child's sleep habits, screen time, and the demands of parenting/caregiving on the respondent.

Section I: About Your Family and Household

Frequency of family meals, the use of tobacco in the home, how the family copes with problems, food adequacy, and adverse childhood experiences. Also, the respondent's perception of their neighborhood (e.g., amenities, safety).

Section J: Child's Caregivers

Demographic information of up to two adults (the respondent and one other person) in the household who are the child's primary caregivers.

Section K: Household Information

Household count, family count, and family income.

2021 Content Changes

Seven variables were added to the 2021 NSCH questionnaires and reported on the public use data files:

- ACE11 ("Has this child EVER been treated or judged unfairly because of a health condition or disability?")
- Four items on health care access during the Coronavirus pandemic:
- COVIDARRANGE ("Have any of this child's regular childcare arrangements been closed or unavailable at any time because of the coronavirus pandemic?")
- COVIDCHECKUPS ("Did this child miss, delay or skip any PREVENTIVE check-ups because of the coronavirus pandemic?")
- VIDEOPHONE ("Has this child had any health care visits by video or phone?")

- VIDEOPHONECOVID (“Were any of this child’s health care visits by video or phone because of the coronavirus pandemic?”)
- VISIONEXAMREC (“Was it recommended that this child see an eye doctor or other eye care provider for an eye examination or additional vision services as a result of the vision screening?”)
- EYECARE (“What care has this child received from the eye doctor?”)
- FRUIT, SUGARDRINK, VEGETABLES (“How many times did this child eat fruit/drink sugary drinks/eat vegetables?”)
- OUTDOORSWKDAY, OUTDOORSWKEND (“ON MOST WEEKDAYS/AN AVERAGE WEEKEND DAY, how much time does this child spend playing outdoors?”)

Responses to these items are reported on NSCH-T1, NSCH-T2, and NSCH-T3 and included in the Topical Public Use File.

Additionally, the questions and response options were updated to reflect gender neutral pronouns. Changes to question wording and response options since 2016 are noted in the NSCH codebook (<https://www.census.gov/data-tools/demo/uccb/nschdict>).

Data Collection Instruments

The data collection design focuses on efforts to increase response rates. Respondents have multiple options to respond to the survey and receive assistance including:

- Web Instrument (English and Spanish)
- Paper Instrument (English and Spanish)
- Telephone Questionnaire Assistance (TQA)
- Email Questionnaire Assistance (EQA)

Web Instrument

The web survey was programmed using the U.S. Census Bureau’s Centurion system for internet data collection. This software presented the questionnaire on a computer screen or other electronic device, e.g., tablet or cellphone. The interview was self-administered by the respondent. The mailed invitation included the survey URL and a unique 8-digit login ID.

Respondents were asked to verify their address. If the respondent answered that the address selected for the sample (and displayed on screen) did not match their own, the survey was concluded and the address was removed from further mailings.

If the listed address matched the respondent’s residence, the case was assigned a PIN that the respondent would need to log back into the survey. Alternatively, the respondent could create a new PIN by correctly answering a security question, which the respondent previously provided during the original PIN creation process.

After setting up the PIN, the respondent reported the number of children (0-17 years of age) that usually resided at that address. If there were no children that usually resided at the address, the survey was concluded and the address removed from further mailings. If there were children that usually resided at the address, the respondent was then directed through the rest of screener questionnaire.

Respondents provided basic demographic and health information for all children in the household (up to 99). There were two hard edits programmed into the web instrument which required the respondent to provide a valid answer before continuing. These answers were necessary for subsampling: child’s first name, initials, or nickname; and child’s age. Respondents were able to skip all other questions and continue the survey.

After the respondent completed the screener questionnaire, the web instrument applied the subsampling methodology. First, the rostered children were sorted by special health care needs (SHCN) status (children with SHCN, then children without SHCN), then by child name (a to z)¹¹, and then by age (youngest to oldest). One of the first four children were randomly selected to be the subject of the topical questionnaire. At this point in the survey process, content from the screener portion of the questionnaire was locked.

The name of the selected child was then prefilled into some topical questions, and the web instrument guided respondents through skip patterns. Some response fields only accepted responses that represented legitimate values; other fields offered a “pick list” of response categories. There were soft edits for some questions that prompted respondents to provide an answer or revise an existing answer, but respondents were able to continue past these edits without changing their answers.

Respondents could review and edit any answers before submitting. Once the survey was submitted, a submission confirmation screen appeared with the date and time of completion. The instrument was then locked and the respondent was only able to view the submission confirmation screen if they logged back in. Submitted responses were saved in the output data file.

Respondents from households without children completed the web instrument in an average of 1 minute, 31 seconds. Respondents from households with children completed the screener portion of the instrument in 5 minutes, 16 seconds; the web topical portion in 30 minutes, 45 seconds; and the entire web instrument in 36 minutes, 1 seconds, on average. Online help screens and text were also available in the instrument to aid respondents.

Table 1. Web Submission Times (in Minutes)

	With Children		No Children	
	Mean	Median	Mean	Median
Screener	5.3	4.1	1.5	0.9
Topical	30.8	26.3		
Total	36.0	31.0	1.5	0.9

Paper Instrument

The second mode of data collection was a two-phase, self-administered mail survey using paper questionnaires. The paper questionnaires were created using Amgraf One Form Plus. They were printed,

¹¹ Children are sorted by name only if there are four or more children in the household. Otherwise, children are sorted only by SHCN and age.

trimmed, and stitched through an in-house print on-demand process using a Docuprint system that allowed personalization to each respondent.

In the first phase of this mode of data collection, paper screener questionnaires were mailed to High Paper addresses with the initial invitation, and to all other addresses (High Web) with the second non-response follow-up mailing.¹² Respondents completed a screener questionnaire to determine if there were any children 17 years of age or younger who usually lived or stayed at the address. Resident children were rostered in the screener instrument. Detailed information was collected for up to four children, while basic information (name, age, sex) was collected for an additional six children.

If the respondent mailed back the screener, it was then processed to determine if eligible children usually resided at the address. Returned forms were processed by iCADE to capture responses through OMR (optical mark recognition), OCR (optical character recognition), and KFI (keying from image). If the respondent answered that the address selected for the sample did not match their own or that there were no children that usually resided at the address, the survey was concluded and the household was removed from further mailings. If the respondent listed children that usually resided at the address, Census Bureau staff applied the subsampling methodology to select one child from the household roster to be the subject of the topical questionnaire.

In the second phase, households that reported eligible children were mailed one of three age-based topical questionnaires requesting more information about the selected child living at the address. To ensure respondents answered the topical questions for the selected child, Docuprint systems printed the selected child's first name, initials, or nickname, age, and sex provided on the screener questionnaire onto the invitation letter and paper questionnaire.

The paper and web instruments were designed to be as similar as possible to minimize the influence of mode on responses. While automatic skips and soft edits could not be implemented in the paper instrument, the questionnaire did include skip instructions within the question wording to mimic the web instrument.

Telephone Questionnaire Assistance (TQA)

The National Processing Center call centers in Jeffersonville, Indiana and Tucson, Arizona provided telephone questionnaire assistance (TQA) for the 2021 NSCH. Respondents could call a toll-free telephone line if they had questions about the survey, wanted to complete the interview over the phone using the web instrument, or submit feedback. All mail content and the web instrument listed this toll-free number.

Interviewers were trained to use the Automated Tracking and Control (ATAC) system to report call-ins using one of the TQA purpose codes seen in Table 2.

¹² More information on the High Web/High Paper group assignments is covered in the Mailout Content and Schedule section.

Table 2. TQA Purpose Codes used in ATAC System

TQA Purpose Codes	Definitions
03	Questionnaire completed - Children in the household
04	Questionnaire completed – No children in the household
02	Refusal to participate
07	Confirmed correct address
08	Confirmed incorrect address
09	Out-of-Scope (vacant, business, not a full-time residence)
10	Spanish questionnaire completed
12	Child moved and/or doesn't live at residence most of the time
20	Questions about incentive
29	Paper questionnaire status
30	Request English paper questionnaire
31	Request Spanish paper questionnaire
32	Trouble completing paper questionnaire
33	Child listed on questionnaire is deceased
51	Centurion issues – PIN and/or LoginID issue
52	Centurion issues – Other
53	Centurion issues – RESET case
60	Questions about the survey (FAQs)
80	None of the above

If any changes were needed to the ATAC TQA instrument based on comments received from interviewers, the survey team coordinated programming updates. All updates to procedures were communicated to the TQA interviewers. Incoming call volumes were also monitored throughout data collection and interviewer schedules were adjusted accordingly.

Email Questionnaire Assistance (EQA)

In addition to the toll-free telephone line, respondents were able to interact with Census Bureau staff via email. An email address (childrenshealth@census.gov) was listed on all invitation letters and on the Centurion login page. Emails were answered by call center staff in Tucson, Arizona. Staff checked the email inbox daily and replied to respondents' messages within 2 business days when possible. Emails were logged in a tracking spreadsheet and cases were assigned purpose codes similar to the TQA purpose codes in Table 2.

EQA agents employed scripted responses for common concerns and questions. These scripts ensured consistent and accurate information. When replying to the messages, agents removed any information in the response email that could be considered personally identifiable (e.g., address, phone number, name).

Spanish Language Translation

The NSCH paper and web instruments were available in both English and Spanish. The Census Bureau reviewed and verified text from the 2020 Spanish-language questionnaires and provided new translations where necessary for the 2021 questionnaires. Respondents could request a Spanish-language questionnaire by calling TQA. Spanish-speaking respondents that called the TQA line were placed in a Spanish-language calling queue; a trained Spanish-language agent then answered any questions or administered the Spanish-language web instrument over the phone. The agent flagged the case if a Spanish paper questionnaire was requested and informed the respondent that a questionnaire would arrive in the mail within three weeks.

If a respondent returned a Spanish-language paper screener questionnaire indicating the presence of children in the household, the Spanish-language topical questionnaire was subsequently mailed to the household. The web instrument included a toggle on every page that allowed respondents to switch between the English and Spanish-language versions of the instrument.

Treatment Groups

Respondent contact strategies and letters were carefully designed to capture the attention of the respondent and pique interest in the subject matter. To increase response and minimize nonresponse bias, the survey sample was divided into treatment and control groups for various experiments. The 2021 NSCH treatments were:

- Screener Cash Incentives
- Topical Cash Incentives
- Mixed Mode (High Paper) vs. Web-Push (High Web)

Screener Cash Incentives

In the initial mailing for screener questionnaires, 90% of the sample received a \$5 bill as an incentive to complete the survey. The other 10% of the sample did not receive an incentive and represented the control group for monitoring the effectiveness of the incentive treatments.

Topical Cash Incentives

Among the households that were mailed a paper topical questionnaire, 70% of households in the first four topical mailing groups and 30% of households in final five topical mailing groups received a \$5 bill as an incentive to complete the survey. The remaining households (30% of the first four groups and 70% of the final five groups) received a \$10 incentive. Total, about 60% of households that were mailed a paper topical questionnaire received a \$10 incentive.

Mixed Mode (High Paper) vs. Web-Push (High Web)

The High Paper treatment group was composed of the 30% of addresses identified as having the highest probability of responding by paper only, and were contacted using a mixed-mode strategy. By random assignment, half of these addresses received a paper screener questionnaire and an invitation to respond by web in the first contact; the other half of the High Paper group received a web invitation in the initial contact and a paper screener questionnaire in the first follow-up mailing. The remaining 70%

of addresses (High Web) were contacted using a web-push strategy. These addresses were mailed only the invitation to respond by web in the first and second contact attempts. More information about the mailout schedule is included in the Data Collection section.

Mailing Contents and Schedule

Data collection for the 2021 NSCH involved a series of mailings and nonresponse follow-up activities, emphasizing questionnaire completion. Mailouts began June 25, 2021 and continued until the survey closeout on January 14, 2022. The approach to data collection and nonresponse follow-up was based on previous project experience and recommendations made by Dillman and colleagues (2009):¹³

- *Invitation letter.* An initial invitation letter was mailed to all potential respondents providing details about the study, a web URL with the login ID for accessing the web version of the survey (which combined the screener and topical into a consolidated instrument), and a toll-free number and email address for individuals to contact if there were questions or comments.
- *Additional mailings.* Subsequent to the first invitation, the Census Bureau sent all remaining non-responding addresses additional invitations. Addresses also received reminder postcards after the first two mailings.

The production mailing schedule for the 2021 NSCH in Table 3 includes screener and topical mailing events. The production mailout schedule consists of up to four screener mailings and two postcard reminders across two groups (High Web and High Paper), and up to four topical mailings and a postcard reminder to each of nine topical mailing groups (A – I below).

Table 3. Production Mailout Schedule

Date	Event
June 25, 2021	Initial Mailing: High Web
June 25, 2021	Initial Mailing: High Paper
July 2, 2021	Pressure Sealed Postcard: High Web
July 2, 2021	Pressure Sealed Postcard: High Paper
July 26, 2021	1st Follow-Up: High Web
July 30, 2021	1st Follow-Up: High Paper
July 30, 2021	2 nd Pressure Sealed Postcard: High Web
August 6, 2021	2 nd Pressure Sealed Postcard: High Paper
August 11, 2021	Topical Mailing 1
August 20, 2021	Topical Pressure Sealed Postcard: Group A
August 20, 2021	2 nd Follow-Up: High Web
August 25, 2021	Topical Mailing 2

¹³ Dillman DA, Smyth JD, Christian LM. 2009. Internet, Mail and Mixed-Mode Surveys: The Tailored Design Method, 3rd edition. Hoboken, NJ: John Wiley & Sons.

Date	Event
September 3, 2021	Topical Pressure Sealed Postcard: Group B
September 3, 2021	2 nd Follow-Up: High Paper
September 9, 2021	Topical Mailing 3
September 17, 2021	Topical Pressure Sealed Postcard: Group C
September 17-24, 2021	3 rd Follow-Up: High Web
September 22, 2021	Topical Mailing 4
October 1, 2021	Topical Pressure Sealed Postcard: Group D
October 6, 2021	3 rd Follow-Up: High Paper
October 7, 2021	Topical Mailing 5
October 15, 2021	Topical Pressure Sealed Postcard: Group E
October 22, 2021	Topical Mailing 6
October 29, 2021	Topical Pressure Sealed Postcard: Group F
November 4, 2021	Topical Mailing 7
November 12, 2021	Topical Pressure Sealed Postcard: Group G
November 19, 2021	Topical Mailing 8
November 26, 2021	Topical Pressure Sealed Postcard: Group H
December 3-4, 2021	Topical Mailing 9
December 10, 2021	Topical Pressure Sealed Postcard: Group I
January 14, 2022	Survey Closeout

Initial Screener Invitation

The initial mailing included the following treatments:

- Screener (\$5) cash incentives
- Mixed-mode (High Paper)

Postcard reminders were mailed one week after initial mailings.

Screener Non-response Follow-up Mailings

The screener non-response follow-up mailings included the following treatments:

- Mixed-mode (High Paper)

All High Paper households received a paper screener questionnaire in the first follow-up mailing.

Postcard reminders were mailed one week after the first follow-up mailing. The screener data collection

strategy included three attempts for non-response follow-up.¹⁴ Addresses remained in their mode assignment (High Paper or High Web) unless a High Web household requested a paper questionnaire before the first follow-up mailing.

Topical Questionnaire

The topical questionnaires were only sent to households that returned a complete paper screener questionnaire, had eligible children in the house, and had not submitted a questionnaire by web. Topical mailings included the following treatments:

- Topical (\$5 or \$10) cash incentives - initial attempt only

There were nine pre-determined mailing dates (1-9) for topical questionnaires. When respondents returned a complete paper screener, they were assigned to the next planned mailing date's initial mailing group (A-I; see Table 4). There were up to three attempts for non-response follow-up depending on the respondent's group assignment. The number of follow-up mailings was constrained by the data collection window, with later groups receiving fewer attempts; groups A-C received three follow-ups, groups D and E received two, groups F and G received one, and groups H and I did not receive follow-up mailings. All topical mailings included a paper topical questionnaire.

Postcard reminders were mailed approximately one week after the initial mailing for that household.

Table 4. Topical Mailings and Topical Mailing Groups

Mailing	Initial	1st Follow-up	2nd Follow-up	3rd Follow-up
Mailing 1	Group A			
Mailing 2	Group B			
Mailing 3	Group C	Group A		
Mailing 4	Group D	Group B		
Mailing 5	Group E	Group C	Group A	
Mailing 6	Group F	Group D	Group B	
Mailing 7	Group G	Group E	Group C	Group A
Mailing 8	Group H	Group F	Group D	Group B
Mailing 9	Group I	Group G	Group E	Group C

¹⁴ Addresses stopped receiving mailings if the residents submitted a web survey, returned a complete paper screener, explicitly refused to participate, or if the address was out-of-scope (i.e., not an occupied residence). The address also received fewer mailings if the USPS determined the address to be undeliverable as addressed.

Response Analysis

Response Rates

Table 5 provides a summary of the survey completion counts. Approximately 108,000 households completed a screener portion of the survey. Of those, 62,010 reported children and are included on the Screener data file.

Complete and sufficient partial topical questionnaires are included on the Topical Public Use File. Of the 62,010 screened households with children, 50,892 returned a complete or sufficient partial topical survey. In 2021, 90.6% of respondents completed the survey using the web instrument and 9.4% of respondents completed the survey using the paper instruments.

Table 5. 2021 Final Dispositions (Unweighted)

Final Disposition	Count
Total Cases	300,000 ^a
Occupied Households (Estimated)	251,000 ^a
Households with Children (Estimated)	154,000 ^a
Completed Screeners	106,000 ^a
Screeners with Children	62,010
Completed Topicals	50,892

^a Rounded to the nearest thousand

For the purposes of calculating response rates, all sampled addresses were assigned screener and topical outcomes codes. These outcomes can be summarized as not eligible, eligible but not complete, complete or eligibility unknown.

For some addresses, there was not sufficient correspondence to determine if the address was eligible to complete the screener or topical questionnaires. These addresses were classified as unresolved. Among these addresses, we estimated the share that were occupied residences using the Household Rate, which is the proportion of resolved addresses that are occupied residences.¹⁵ We also estimated the Child Rate, which is the share of those households that include children, based on the proportion of households that have children by state and stratum in the 2019 American Community Survey (ACS). The product of the Household Rate and Child Rate is the Eligibility Rate (e), the estimated proportion of unresolved addresses that are households with children. Using this approach, we estimated that 87% (weighted) of unresolved addresses were households, and 40% (weighted) of those were households with children.

¹⁵ Specifically, we used the midpoint between the Household Rate including undeliverable addresses (the proportion of all resolved addresses that are occupied residences) and the Household Rate excluding undeliverable addresses (UAA) by state and stratum. Because UAAs are identified by the U. S. Postal Service, it is assumed that UAAs are identified at a higher rate than other noneligible addresses (businesses, vacant residences, etc.) that must be self-identified. The midpoint assumes that there are some UAAs still unresolved but at a lower rate than they appear among the resolved addresses.

$$e = \text{Household Rate} * \text{Child Rate}$$

Three different response rates were calculated based on the estimated proportion of eligible addresses that completed the screener and topical questionnaires. Definitions of completion and the calculation of these three response rates are detailed below.

Screener Completion Rate

The Screener Completion Rate (SCR) is the estimated proportion of households (occupied residences) that completed a screener. A completed screener had to 1) be returned from a sampled address, and 2) indicate that there were no children present or provide a valid age for at least one child. The denominator includes both screened households and the number of unresolved addresses that are estimated to be households.

$$SCR = \frac{\text{Completed Screeners}}{\text{Screened HHs} + (\text{Unresolved Addresses} * \text{Household Rate})}$$

Topical Completion Rate

The Topical Completion Rate (TCR) is the estimated proportion of households with children that returned a topical questionnaire, either complete or sufficient partial. Completed topical questionnaires have valid answers for at least 40 of 50 test questions. Also, at least one item in Section K (family income, household and family count) must be completed, or the respondent submitted the topical portion of the web instrument. Sufficient partial topical questionnaires have valid answers for at least 25 of 50 test questions. Also, at least one item in Section H or beyond must be completed, or the respondent submitted the topical portion of the web instrument. The denominator includes both screened households with children and the number of unresolved addresses that are estimated to be households with children (Unresolved Addresses * e).

$$TCR = \frac{\text{Completed Topicals}}{\text{Screened HHs with Children} + (\text{Unresolved Addresses} * e)}$$

Interview Completion Rate and Overall Response Rate

The Interview Completion Rate (ICR) and Overall Response Rate (ORR) account for the multi-stage design of the NSCH. They are the products of two (for ICR) or three (for ORR) response rate metrics that are each consistent with the American Association for Public Opinion Research (AAPOR) standards.¹⁶

The ICR is the probability that a household that initiates the survey will complete the screener and, if applicable, the topical portions of the survey. It is calculated as the product of the Screener Conversion Rate (the proportion of screeners that are completed) and the Topical Conversion Rate (the proportion of households that complete a detailed topical questionnaire after reporting children on a completed screener).

$$ICR = \text{Screener Conversion Rate} * \text{Topical Conversion Rate}$$

¹⁶ The American Association for Public Opinion Research. 2016. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. 9th edition. AAPOR.

The ORR is the probability that an address progresses through the three major stages of survey completion – resolution, screener, and topical questionnaire.

$$ORR = \text{Resolution Rate} * \text{Screener Conversion Rate} * \text{Topical Conversion Rate}$$

The Resolution Rate is the proportion of addresses that are resolved, i.e., returned sufficient information to determine if the address represents a household.¹⁷ In 2021, the weighted Resolution Rate was 50.7%.

$$\text{Resolution Rate} = \frac{\text{Resolved Addresses}}{\text{Total Addresses}}$$

Table 6 lists the weighted rate for each of the four response metrics discussed above. A breakdown of the response rates by state is provided in Attachment D.

Table 6. 2021 NSCH Weighted Response Rates

Metric	Rate
Screener Completion Rate	45.4%
Topical Completion Rate	32.3%
Interview Completion Rate	79.5%
Overall Response Rate	40.3%

Item-Level Response

The item-level response rate is the proportion of item-eligible respondents that provided a valid answer to a particular item. Many items were applicable to a subset of survey respondents only; for example, some questions were applicable to children in a specific age range. In that case, the denominator for the item-level response rate is the count of children in the eligible age range, and the numerator is the count of those children with valid responses.

In some cases, it is uncertain if the child was eligible for an item due to nonresponse on a preceding item. For example, before asking about the severity of certain conditions, we asked if the child currently had the condition. The severity item was applicable if the child currently had the condition, and it was not applicable if the child did not currently have the condition. If the respondent chose to skip the current condition filter item, we cannot know definitively if the severity item was applicable or not.

We account for this situation in the item-level response rate by assigning eligibility to cases with unknown eligibility equal to the proportion of cases that were eligible when eligibility was known. For example, if 10% of respondents reported that the child did have the condition currently, and so were eligible for the severity follow-up question, the denominator for the severity item response rate becomes

$$\# \text{ Eligible} + (\# \text{ Eligibility Unknown} * .1)$$

¹⁷ An address is resolved as a household when the household begins the screener portion of the survey.

Across all survey items, more than 98% of response opportunities produced a valid response. Items that require a write-in response, that require respondents to follow a skip pattern, and are near the end of the survey tend to have higher nonresponse. Table 7 lists the 35 variables with the lowest item-level response rates. The list predominantly reflects items that are at the end of a skip pattern and are on-path for few respondents (e.g., CYSTFIB_DESC), items that require a write-in response (e.g., K2Q35A_1_YEARS), and items near the end of the survey (e.g., A2 items).¹⁸

Table 7. Lowest Item-Level Response Rates

Variable	Description	Response Rate	On-Path (%)
CYSTFIB_DESC	Cystic Fibrosis Severity Description	84.3%	0.1%
VIDEOPHONECOVID	Virtual Health Care Visits Due to COVID	85.6%	24.9%
HEMOPHILIA	Blood Disorder Hemophilia	86.5%	0.5%
SICKLECELL	Blood Disorder Sickle Cell	87.7%	0.5%
THALASSEMIA	Blood Disorder Thalassemia	89.0%	0.5%
K2Q35A_1_YEARS	Autism ASD - First Told Age in Years	91.6%	3.0%
BLOOD_OTHER	Blood Disorder Other (includes Hemophilia)	91.9%	0.5%
HCEXTENT	Health Affected Ability - Extent	93.4%	29.1%
BLOOD_DESC	Blood Disorder Severity Description	94.0%	0.4%
A2_LIVEUSA	Adult 2 - Come to Live in the United States (Year)	94.2%	13.7%
A2_BORN	Adult 2 - Where Born	95.2%	85.3%
SPCSERVMO	Received Special Services - Age in Months (use with K4Q37)	95.6%	17.1%
K4Q37	Received Special Services - Age in Years	95.6%	17.1%
A2_DEPLSTAT	Adult 2 - Deployment Status	95.6%	5.5%
ACE12	Treated Unfairly Because of their Sexual Orientation or Gender Identity	95.7%	59.3%
A2_MARITAL	Adult 2 - Marital Status	95.8%	85.3%
COVIDARRANGE	Closed Childcare or Daycare	95.8%	68.2%
ACE10	Child Experienced - Treated Unfairly Because of Race	95.9%	100.0%
A2_PHYSHEALTH	Adult 2 - Physical Health	95.9%	85.3%
K12Q01_G	Reason Not Covered - Other	95.9%	5.4%
K12Q01_F	Reason Not Covered - Application/Renewal Problems	95.9%	5.4%
K12Q01_E	Reason Not Covered - Inadequate Providers	95.9%	5.4%
K12Q01_D	Reason Not Covered - Inadequate Benefits	95.9%	5.4%

¹⁸ This table does not include the six poverty status implicates (FPL_I1-FPL_I6). Values for these items are derived from several survey items, and partial responses are used to inform the multiple imputation. For comparison, 19.7% of respondents do not provide sufficient information to deduce the poverty status from survey responses alone.

K12Q01_C	Reason Not Covered - Unaffordable	95.9%	5.4%
K12Q01_B	Reason Not Covered - Cancellation Overdue Premiums	95.9%	5.4%
K12Q01_A	Reason Not Covered - Change in Employer/Employment	95.9%	5.4%
A2_EMPLOYED	Adult 2 - Current Employment Status	95.9%	85.3%
ACE8	Child Experienced - Lived with Mentally Ill	95.9%	100.0%
A2_MENTHEALTH	Adult 2 - Mental or Emotional Health	95.9%	85.3%
ACE7	Child Experienced - Victim of Violence	95.9%	100.0%
A2_SEX	Adult 2 - Sex	96.0%	85.3%
A1_DEPLSTAT	Adult 1 - Deployment Status	96.0%	4.4%
ACE6	Child Experienced - Adults Slap, Hit, Kick, Punch Others	96.0%	100.0%
STRENGTHS	Facing Problems - How Often Draw on Strengths	96.0%	100.0%
ACE9	Lived with Person with Alcohol/Drug Problem	96.0%	100.0%

Treatment Groups and Response

This section reviews response patterns based on the treatment group assignments:

- Screener (\$5) Cash Incentives
- Topical (\$5 or \$10) Cash Incentives
- “Low-Cost” High Paper Treatment

Screener Cash Incentives

In the initial mailing for screener questionnaires, 90% of the sample received a \$5 bill as an incentive to complete the survey.

The unconditional cash incentives are included with the initial invitation to encourage households to respond. The results of the intervention are reported in Table 8. Eligible households that received a \$5 incentive were more likely to complete the Screener and Topical questionnaires than households that received no incentive. Extrapolating from the results below and the estimate in Table 5 that the 2021 sample included 154,000 households with children, we estimate that 8,800 fewer households would have completed the topical questionnaire if we did not use screener incentives.

Table 8. Data Collection Costs and Completed Questionnaires by Screener Incentive

Screener Cash Incentive Group	Average Cost per Completed Screener	Percent of Eligible Households that Completed a Screener	Average Cost per Completed Topical	Percent of Eligible Households that Completed a Topical
Total	\$26.79	42.1%	\$59.52	33.2%
No Incentive	\$17.82	36.3%	\$43.23	27.4%
\$5 Incentive	\$27.63	42.7%	\$60.98	33.8%

Topical Cash Incentive

For households who were mailed their first paper topical questionnaire, roughly 60% received a \$10 bill as an incentive. The other 40% received a \$5 incentive. The distribution of incentives depended on mailing group. The early mailing groups (A - D), that tend to have a higher response rate than the later mail groups, received a smaller share of \$10 incentives (30% versus 70%). .

Table 9. Data Collections Costs and Returns by Topical Incentive

Incentive Group	Average Cost per Completed Topical	Completion Rate*
Groups A - D	\$47.43	60.9%
\$5 Incentive	\$45.52	60.4%
\$10 Incentive	\$51.79	62.1%
Groups E - I	\$63.39	46.3%
\$5 Incentive	\$58.78	44.2%
\$10 Incentive	\$65.24	47.2%

*Percent of households that were mailed a paper topical invitation that subsequently completed a topical interview.

Households that received a \$10 topical incentive were more likely to complete the topical questionnaire than households that received the \$5 topical incentive (see Table 9). The difference varied by topical mail group; the larger incentive increased the probability of response for the later mail groups by 6.9%, but only increased response by 2.7% for the early mail groups.

Incentives and Non-Response Bias

In past cycles of the NSCH, cash incentives were relatively more effective among groups that were otherwise less likely to respond.¹⁹ The incentives did not have the same effect in 2020, when instead the cash incentive significantly increased screener and topical response for all groups. In 2021, incentive performance was again consistent with our experience prior to 2020.

Table 10 shows the incentive effect as the ratio of the probability of screener response with and without the incentive by education, race and poverty/income. The cash incentive significantly increased response for all groups, and the incentive effect was significantly larger for less educated households, Black households (compared to White households), and households in poverty. Because less educated, Black and poor households are generally less likely to respond to the survey, incentives may reduce nonresponse bias.

¹⁹ NSCH 2019 Methodology Report (<https://www2.census.gov/programs-surveys/nsch/technical-documentation/methodology/2019-NSCH-Methodology-Report.pdf>)

Table 10. Response Odds Ratios (Incentive versus No Incentive) by Education, Race, and Income

Education	Screener	
	P(\$5 incentive) /	P(control)
High School or less	118.4%	*
College or more	116.6%	*
HS vs. College	+1.7%	†
Race		
Black alone	124.9%	*
White alone	116.4%	*
Black vs. White	+7.9%	†
Poverty Status		
Poverty	119.4%	*
Income > \$100k	116.1%	*
Poverty vs. Income > \$100k	+3.3%	†

* H₀: P(incentive)/P(control) ≤ 1, p<0.05; † H₀: Difference ≤ 0, p<0.05

“Low-Cost” High Paper Treatment

The mixed-mode data collection strategy used with the High Paper addresses encourages more households to respond by mail, which has two disadvantages compared with web response. 1) Mail-out/mail-back data collection is more costly than web. 2) Households with children that respond to the screener questionnaire online are more likely to complete the topical questionnaire than those that mail back a paper screener questionnaire. On the other hand, providing a paper questionnaire early in the data collection process facilitates response from households that prefer to respond by mail over web. In 2021, we tested a “low-cost” High Paper treatment as a compromise between the traditional High Paper and High Web treatments. Specifically, these households received a web-only invitation in the initial contact and a web invite along with a paper questionnaire in the first nonresponse follow-up. Half of households from the High Paper group were randomly assigned to the low-cost treatment.

Comparing the low-cost group to the traditional High Paper group, the results were largely as anticipated. Average data collection costs were \$1.65 lower for the low-cost group, roughly equal to the cost of the paper questionnaire that was not included with the initial invitation. On the other hand, screener response was 2.6 percentage points lower for this group than the traditional High Paper group. Topical response was not meaningfully different between the two groups.

Data Processing

Data were processed and edited to ensure data quality and respondent confidentiality.

Unduplication

All nonresponding households were offered two modes, web and paper, for completing the survey. In some cases, respondents utilized both options. In these cases, we selected one response, web or paper, to include in the data file. We chose the response to include based on the type of return and the level of completeness. Completed web returns were always chosen over completed paper returns. However, completed paper returns were chosen over partial web survey returns. The web/paper unduplication hierarchy is detailed in Table 11.

Table 11. Unduplication Criteria for both Web and Paper Returns

Order Chosen	Type of Return
1	Completed web survey - Household with children
2	Completed paper screener and topical
3	Completed web survey - Household w/o children
4	Completed paper screener - Household w/o children
5	Partially completed web survey
6	Out of scope paper return
7	Refusal paper return, Hard Refusal
8	Incomplete, Duplicate
9	Blank, Soft Refusal
10	Deceased
11	Undeliverable address (UAA) with address correction – mail forwarded, UAA with address correction
12	UAAs, Forwarding Order Expired, Moved out of U.S.
13	Default
14	Blank form

Multiple follow-up mailings included paper questionnaires, so it was also possible that respondents received and returned more than one questionnaire. In these cases, one return was selected to represent that case in the data file. The paper/paper unduplication hierarchy is detailed in Table 12.

Table 12. Unduplication Criteria for Two Paper Returns

Order Chosen	Type of Return
1	Completed paper screener/topical - Household with children
2	Completed paper screener - Household w/o children
3	Out of scope paper return
4	Refusal paper return, Hard Refusal
5	Incomplete, Duplicate
6	Blank, Soft Refusal
7	Deceased
8	UAA with address correction – mail forwarded, UAA with address correction
9	UAAs, Forwarding Order Expired, Moved out of U.S.
10	Default
11	Blank form

Paper to Web Standardization

Responses were standardized across web and paper so they could be appended in a single data file. Although survey questions had the same valid values for the paper and web instruments, sometimes the values output for the paper questionnaire did not match the output from the web survey instrument. For instance, any questions that included a list of checkboxes where the respondent was instructed to “Mark (X) ONE box” differed between paper and web. The web instrument had the ability to prevent the selection of more than one checkbox via a radio button, whereas a paper respondent could mark more than one box even if the question explicitly said not to. Since all data from the paper instruments is captured for processing, each of the response option checkboxes have their own associated output variable. Therefore, prior to appending web and paper responses into a single data file, paper responses were reformatted to the proper valid values.

Data Edits

The 2021 NSCH raw output was processed to manage inconsistent and invalid responses in nine sequential steps:

- *Stop Process Edit.* A case is removed from the data file if the case fails address verification (the respondent indicates that their address does not match the address on file), the respondent indicates that there are no children in the household, or the respondent does not complete a screener for a household with children. The cases are not eligible to be included on a NSCH data file, so are removed from processing.
- *Not in Universe Edit.* An item is not in universe if it is not included in the instrument the respondent received. Some items are unique to web or paper, and others are specific to a

version of the topical instrument, T1, T2, or T3. The value for an item that is not in universe is set to '.N'.

- *Range Edit.* If a value falls outside the bounds of a defined minimum and maximum for that item, the value is replaced with an indicator that the response is missing. The minimum and maximum are selected to represent a reasonable range of possible responses to the item.
- *Backfill Edit.* The backfill edit imputes values to some items based on responses to subsequent items that necessarily indicate the correct response to the edited item. Backfill edits apply almost exclusively to paper questionnaires, which cannot prevent a respondent from skipping a root item but answering follow-up questions. For example, INCWAGES is a binary item that filters respondents on whether the family did (INCWAGES=1) or did not (INCWAGES=2) receive wage or salary income. If a respondent does not answer INCWAGES, but provides a valid and non-zero value for INCWAGES_AMT, the dollar amount of wage and salary income, then it is necessarily correct that INCWAGES=1.
- *Yes/No Edit.* The NSCH includes several series that ask respondents to select all applicable items from a list. These series may or may not allow the respondent to answer in the negative, indicating that the item is not applicable. In most cases, if a respondent answers in the affirmative (=1) to at least one item in the series, it is assumed that all other items in the series do not apply (=2) unless otherwise noted. If a respondent is only able to respond in the affirmative, and the items in the series are not comprehensive (e.g., they do not include an "Other" option), then it is assumed that all unanswered items do not apply (=2) without imposing the requirement that at least one item is answered in the affirmative.
- *Consistency Edit.* If responses to two items in the survey are fundamentally inconsistent, one response is maintained and the other is removed and changed to missing. Most consistency edits require that a child does not experience a life event at an age greater than their current age. Because the instrument generally trends from more general, fundamental information to more specific, priority is given to the item that appears first in the instrument.
- *Legitimate Skip Edit.* Unlike the 'Not in Universe Edit', the legitimate skip edit applies to items that are on the respondent's instrument, but not on path. The value for an item that is in universe but not on path is set to '.L'.
- *Missing in Error Edit.* If an item is in universe (does not equal .N), is on path (does not equal .L), but does not hold a valid value, that item is missing in error, identified as '.M'.
- *Disclosure Edit.* Some survey responses, if published, could compromise a respondent's confidentiality. Disclosure edits involve removing entire items (e.g., child's name) or suppressing rare or unique values (e.g., top codes on the family poverty ratio). Census disclosure avoidance

standards make reference to weighted and unweighted cell counts (i.e., the number of children with a characteristic or set of characteristics), the size of the underlying population (e.g., the number of children in Kentucky Metropolitan Statistical Areas), and the existence of outside data sources that could be matched to the NSCH (e.g., a registry of children diagnosed with Cerebral Palsy).

Edits were applied in two stages. In the first stage, edits for screener items were applied to completed screeners with children. When these edits were completed, cases that did not return a completed topical were removed from edits, and the second stage edits to topical items were applied.

Recoded and Standardized Variables

Standardized Variables

Several questions in the 2021 NSCH allowed respondents to provide an answer using more than one unit (e.g., years and months) and to choose from two systems of units (e.g., imperial or metric). In these cases, we provide standardized variables that convert responses across units and systems to a single unit. See Table 13 for a list and description of these variables.

Table 13. Standardized Variables

Variable	Description	Units
BIRTHWT_OZ_S	Child birth weight	Ounces
BREASTFEDEND_DAY_S	Stopped breastfeeding	Days
BREASTFEDEND_WK_S	Stopped breastfeeding	Weeks
BREASTFEDEND_MO_S	Stopped breastfeeding	Months
FRSTFORMULA_DAY_S	First fed formula	Days
FRSTFORMULA_WK_S	First fed formula	Weeks
FRSTFORMULA_MO_S	First fed formula	Months
FRSTSOLIDS_DAY_S	First fed solids	Days
FRSTSOLIDS_WK_S	First fed solids	Weeks
FRSTSOLIDS_MO_S	First fed solids	Months

Derived and Recoded Variables

A number of variables on the public use data files are derived from a set of items on the survey or a recoded version of a single item. These variables are listed in Table 14.

Table 14. Derived and Recoded Variables

Variable	Description	Derived from
AGEPOS4	Birth position of the selected child relative to other children in household	C_AGE_YEARS C_AGE_MONTHS
TOTMALE	Count of male children in household	C_SEX
TOTFEMALE	Count of female children in household	C_SEX

Variable	Description	Derived from
C_CSHCN	Special Health Care Needs (SHCN) status	C_K2Q10 - C_K2Q23
SC_CSHCN	SHCN status of selected child	C_CSHCN
TOTCSHCN	Count of children with SHCN	CSHCN
TOTNONSHCN	Count of children that do not have SHCN	C_K2Q10 - C_K2Q23
TOTAGE_0_5	Count of children 0 to 5 years old in household	C_AGE_YEARS
TOTAGE_6_11	Count of children 6 to 11 years old in household	C_AGE_YEARS
TOTAGE_12_17	Count of children 12 to 17 years old in household	C_AGE_YEARS
SC_AGE_LT4	Age of selected child (less than 4 months)	SC_AGE_YEARS SC_AGE_MONTHS
SC_AGE_LT6	Age of selected child (less than 6 months)	SC_AGE_YEARS SC_AGE_MONTHS
SC_AGE_LT9	Age of selected child (less than 9 months)	SC_AGE_YEARS SC_AGE_MONTHS
SC_AGE_LT10	Age of selected child (less than 10 months)	SC_AGE_YEARS SC_AGE_MONTHS
C_RACER	Race of child	C_RACE_R
C_RACEASIA	Asian race category is included for the following states: CA, HI, MA, MD, MN, NJ, NV, NY, VA, WA	C_RACE_R
C_RACEAIAN	American Indian/Alaska Native race category is included for the following states: AK, AZ, NM, MT, ND, OK, SD	C_RACE_R
C_HISPANIC_R	Hispanic origin	C_HISPANIC
SC_RACER	Race of selected child	SC_RACE_R
SC_RACEASIA	Asian race category is included for the following states: CA, HI, MA, MD, MN, NJ, NV, NY, VA, WA (Selected Child)	SC_RACE_R
SC_RACEAIAN	American Indian/Alaska Native race category is included for the following states: AK, AZ, NM, MT, ND, OK, SD (Selected Child)	SC_RACE_R
SC_HISPANIC_R	Hispanic origin of selected child	SC_HISPANIC
HOUSE_GEN	Parental nativity	BORNUSA A1_RELATION A1_BORN A2_RELATION A2_BORN

Variable	Description	Derived from
FAMILY_R	Family structure	A1_RELATION A2_RELATION A1_MARITAL A2_MARITAL A1_SEX A2_SEX
CURRINS	Current health insurance coverage status	K3Q04_R CURRCOV K12Q03, K12Q04, K12Q12, TRICARE, HCCOVOTH, K11Q03R
INSTYPE	Type of insurance	CURRINS K12Q03, K12Q04, K12Q12, TRICARE, HCCOVOTH, K11Q03R
INSGAP	Health insurance coverage over the past 12 months	K3Q04_R, CURRINS
FPL_I1-FPL_I6	Family poverty ratio	FAMCOUNT TOTINCOME
HIGRADE	Highest level of education for reported adults (three categories)	A1_GRADE A2_GRADE
HIGRADE_TVIS	Highest level of education for reported adults (four categories)	A1_GRADE A2_GRADE
BIRTHWT	Birth weight status	BIRTHWT_OZ_S
BIRTHWT_L	Low birth weight (<2500g)	BIRTHWT_OZ_S
BIRTHWT_VL	Very low birth weight (<1500g)	BIRTHWT_OZ_S
BMICLASS	Body Mass Index	WEIGHT_* HEIGHT_*

Specifications of Select Derived Variables

The 2021 NSCH reports several derived variables that include information on the child's family status. This includes Family Poverty Ratio (FPL), Household Nativity (HOUSE_GEN), and Family Structure (FAMILY_R).

- Family Poverty Ratio (FPL)*** - The family poverty ratio is calculated as the ratio of total family income to the family poverty threshold and reported as a rounded percentage. Respondents reported total family income in item K4 on the paper instrument: "The following question is about your 2020 income. Think about your total combined family income IN THE LAST CALENDAR YEAR for all members of the family. What is that amount before taxes?" Additional text instructed respondents to include all money incomes, for example, social security, dividends, and child support. Responses to K4 were edited for consistency against answers in K3,

a series of questions about specific sources of income. Finally, missing or invalid responses were replaced with multiply imputed values.

The family poverty threshold is derived from the Census Bureau's poverty thresholds. Thresholds vary by family size and the number of related children under 18 years old. They do not vary across geographies. Family size was reported in K2 of the paper instrument. Missing or invalid values were imputed. The number of related children was determined by the number of children reported in the screener.

To protect the confidentiality of respondents, only FPL is reported in the Public Use File; total family income and the family poverty threshold are not included. Further, FPL is top and bottom coded. Reported values range from 50 (total family income is 50% of the family poverty threshold) to 400 (total family income is 400% of the family poverty threshold). Values beyond this range are reported as 50 or 400, respectively.

- *Household Nativity (HOUSE_GEN)* - Household nativity is determined by the birth location of the child (BORNUSA) and parents (A1_BORN and A2_BORN). If the child was born outside of the U.S. and all reported parents were born outside of the U.S., the household is reported as a 1st generation household. Second generation households have members born both inside and outside of the U.S. For example, the child was born in the U.S. and at least one parent was born outside of the U.S., or the child was born outside of the U.S. and one of two parents was born in the U.S. Finally, in 3rd+ generation households, all parents were born in the U.S. The fourth category, "Other", captures households with insufficient information about the nativity of the parents.
- *Family Structure (FAMILY_R)* - A family structure variable uses the reported information on the child's primary caregivers to organize households into common types. Notably, the NSCH collects information on only two adults in the household and requires only that the two adults be primary caregivers of the child. As a result, in multigenerational households, this can mean that a biological, adoptive, or stepparent is not reported.

Further, respondents do not report their relationship to other adult members of the household, only to the child; consequently, we may know that the two reported adults are married, but we do not know if they are married to each other. Instead of making assumptions about the relationship of the reported adults with each other, the family structure variable depends only on the number of adults, their relationship to the child, and their individual marital statuses. For example, a reported value of 1 for FAMILY means that the two reported adults are biological/adoptive parents of the child and they are currently married; one may assume that they are married to each other, but in some cases that will not be true.

Two family structure categories (FAMILY_R=5 and 6) are also defined by the sex of the respondent. In these cases, it is specified that the responding caregiver is female (5) or male (6) and that no other parents (biological, adoptive, or step) are in the household.

The 2021 NSCH reports several variables that include information on the child's health insurance status and insurance type. We strongly recommend that data users interested in current health insurance status and insurance type use the derived variables CURRINS (Currently Insured), INSGAP (Gaps in Coverage), and INSTYPE (Insurance Type) in their analyses.

- *Currently Covered (CURRINS)* - CURRINS is derived primarily from the respondent-reported values in K3Q04_R (Health Insurance Coverage – Past 12 Months) and CURRCOV (Health Insurance Coverage – Currently Covered). We indicate that the child is currently insured (CURRINS=1) if the respondent reported that the child had coverage for all of the last 12 months (K3Q04_R=1) or reported that the child is currently covered (CURRCOV=1), but with an important caveat. If the respondent reported that the child is currently insured but reported only Indian Health Service or health care sharing ministry as the type of coverage, we indicate that the child does not have current insurance coverage (CURRINS=2). Consequently, a respondent may report that a child is insured, but we consider that the child is not insured.
- *Gaps in Coverage (INSGAP)* - INSGAP is derived primarily from the respondent reported values in K3Q04_R (Health Insurance Coverage – Past 12 Months) and CURRCOV (Health Insurance Coverage – Currently Covered). We indicate that the child had consistent coverage (INSGAP=1) if the respondent reported that the child had coverage for all of the last 12 months (K3Q04_R=1) but with an important caveat. If the respondent reported that the child is currently insured but reported only Indian Health Service or health care sharing ministry as the type of coverage, we indicate that information as to the consistency of the child's coverage is missing (INSGAP=.M).
- *Insurance Type (INSTYPE)* - INSTYPE is derived from CURRINS (Currently Insured) and respondent answers to questions on the coverage type: K12Q03 (Current/Formal Employer or Union), K12Q04 (Directly Purchased), K12Q12 (Government Assistance Plan), TRICARE (TRICARE or other military health care), K11Q03 (Indian Health Service), and HCCOVOTH_WRITEIN (Other Type, Write-in). Any insurance reported as coming from an employer or union, directly purchased, TRICARE or other military health care, or the Affordable Care Act is considered private. Coverage from any government assistance plan is considered public. Both the private and public coverage categories reflect a single reported source of coverage; a combined category for children with both public and private coverage is also included.

In addition, Health Insurance write-in (HCCOVOTH_WRITEIN) responses were back-coded to flag public and private insurance types, religious health care sharing ministry, and Indian Health Service coverage. These flags were used in the derivation of CURRINS and INSTYPE. To protect respondent confidentiality, answers to HCCOVOTH_WRITEIN are not reported in the Public Use File.

Missing Values and Imputation

For most variables in the public data files, missing values are coded to identify the type of missing data. These include

- (.L) Legitimate Skip – The item is not applicable to the respondent, as determined by a previous answer to a root question.
- (.M) Missing in Error – The value is missing due to respondent or system errors, or the respondent did not provide a valid answer.
- (.N) Not in Universe – The item was not included on the respondent’s age-appropriate version of the topical questionnaire.
- (.D) Suppressed for Confidentiality – The value is suppressed in order to protect respondent confidentiality.

However, variables use for raking during weighting procedures require imputation. Table 15 lists the 2021 variables that are imputed and includes the imputation flag variables to indicate records with imputed values. Tenure, sex, race, and Hispanic origin were imputed using hot-deck imputation. Adult 1 education, household size, and poverty ratio were imputed using sequential regression imputation methods.²⁰

Table 15. Imputed Variables and Their Imputation Flags

Variable	Missing Rate	Imputation Flag Variable
Household tenure (TENURE)	1.23%	Flag for Household Tenure (TENURE_IF)
Child’s sex (C_SEX)	0.16%	Flag for child’s sex (C_SEX_IF)
Child’s race (C_RACE_R)	0.61%	Flag for child’s race (C_RACE_R_IF)
Child’s Hispanic origin (C_HISPANIC_R)	0.51%	Flag for child’s Hispanic origin (C_HISPANIC_R_IF)
Selected child’s sex (SC_SEX)	0.07%	Flag for selected child’s sex (SC_SEX_IF)
Selected child’s race (SC_RACE_R)	0.36%	Flag for selected child’s race (SC_RACE_R_IF)
Selected child’s Hispanic origin (SC_HISPANIC_R)	0.29%	Flag for selected child’s Hispanic origin (SC_HISPANIC_R_IF)
Adult 1’s highest completed year of school (A1_GRADE)	3.39%	Flag for adult 1’s highest completed year of school (A1_GRADE_IF)
Household size (HHCOUNT)	3.10%	Flag for household size (HHCOUNT_IF)
Family poverty ratio (FPL)	18.78%	Flag for family poverty ratio (FPL_IF)

²⁰ For more information on data analysis using imputed values, see <https://www2.census.gov/programs-surveys/nsch/technical-documentation/methodology/NSCH-Analysis-with-Imputed-Data-Guide.pdf>

Multiple Imputation

Using sequential regression imputation methods, FPL is multiply imputed and contains six versions or implicates. The public use file includes all six imputed values for FPL [FPL_I1-FPL_I6]. The primary motivation for the multiple imputation is to allow interested researchers to appropriately account for uncertainty in estimates using FPL that is hidden when using a single implicate.²¹ FPL input includes imputed values for family income (not included in the public use file) and number of people that are family members (FAMCOUNT). An estimated family count (FAMCOUNT) was derived from HHCOUNT and other household information when FAMCOUNT was not reported by the household. The imputation was executed by sequential regression modeling imputation²² using IVEWare.²³

Suppressed Variables

A number of variables had range caps or suppressed values to protect respondent confidentiality consistent with U.S. Census Bureau protocols. For example, a reported value must represent at least 10,000 children (weighted estimate). These variables are listed in Table 16.

Table 16. Suppressed Variables

Variable	Description	Valid Values
TOTKIDS_R	Number of children living in the household	1 = 1 2 = 2 3 = 3 4 = 4+
MOMAGE	Age of mother when child was born	18 = 18 years or younger 45 = 45 years or older
K2Q35A_1_YEARS	Age of child when first diagnosed with autism	1 = 1 year or younger 15 = 15 years or older
BIRTHWT_OZ_S	Birth weight	72 = 72 oz. or less 155 = 155 oz. or more
K11Q43R	Number of time the child has moved to a new address	15 = 15 or more times
A1_AGE	Age of Adult 1	75 = 75 years or older
A2_AGE	Age of Adult 2	75 = 75 years or older
A1_LIVEUSA	When Adult 1 came to live in the U.S.	1970 = Before or in 1970
A2_LIVEUSA	When Adult 2 came to live in the U.S.	1970 = Before or in 1970
BREASTFEDEND_DAY_S	Stopped breastfeeding, age in days	Suppressed if > 5

²¹ Schaefer JL, Graham JW. 2002. "Missing Data: Our View of State of the Art". *Psychological Methods*, 7(2): 147-77.

²² Raghunathan TE, Lepkowski JM, Hoewyk JV, Solenberger PW. 2001. "A Multivariate Technique for Multiply Imputing Missing Values using a Sequence of Regression Models". *Survey Methodology*, 27: 85-95.

²³ Raghunathan TE, Solenberger PW, Hoewyk JV. 2016. IVEware: Imputation and Variance Estimation Software User's Guide (Version 0.3). Ann Arbor, MI: Institute for Social Research, University of Michigan.

Variable	Description	Valid Values
BREASTFEDEND_WK_S	Stopped breastfeeding, age in weeks	Suppressed if > 8
BREASTFEDEND_MO_S	Stopped breastfeeding, age in months	29 = 29 or more
FRSTFORMULA_DAY_S	First fed formula, age in days	Suppressed if > 6
FRSTFORMULA_WK_S	First fed formula, age in weeks	Suppressed if > 8
FRSTFORMULA_MO_S	First fed formula, age in months	12 = 12 or more
FRSTSOLIDS_DAY_S	First fed solids, age in days	Suppressed if > 1
FRSTSOLIDS_WK_S	First fed solids, age in weeks	Suppressed if > 4
FRSTSOLIDS_MO_S	First fed solids, age in months	15 = 15 or more
FPL	Family poverty ratio	50 = 50% or less 400 = 400% or more
FAMCOUNT	Family Count	8 = 8 or more
HHCOUNT	Household Count	10 = 10 or more
K4Q37	Received Special Services - Age in Years	15 = 15 or more
SESPLANYR	Special Education Plan - Age in Years	16 = 16 or more
SESPLANMO	Special Education Plan – Age in Months	Suppressed if SESPLANYR > 2

Geography Variables

The 2021 NSCH includes four geography variables on the Public Use File

- FIPSST (State of Residence)
- CBSAFP_YN (Core-Based Statistical Area Status)
- METRO_YN (Metropolitan Statistical Area Status)
- MPC_YN (Metropolitan Principal City Status)

Table 17 provides a general description of the geography variables and their valid values. To protect respondent confidentiality, CBSAFP_YN, METRO_YN, and MPC_YN are not reported in some states. If a variable or intersection of variables could be used to identify a geographic area within a state with a population under 250,000, reported values for that variable were replaced with ".D", indicating "Suppressed for Confidentiality".

Table 17. List of Geography Variables

Variable	Description	Valid Values
FIPSST	State of Residence	[FIPS code]
CBSAFP_YN	Core Based Statistical Area (CBSA): County or counties associated with at least one core (urbanized area or urban cluster) of at least 10,000 population, plus adjacent counties having a high degree of social and economic integration with the core as measured through commuting ties.	.D = Suppressed for confidentiality 1 = In a CBSA 2 = Not in a CBSA

Variable	Description	Valid Values
METRO_YN	Metropolitan Statistical Area (MSA): County or counties associated with at least one urbanized area of at least 50,000 population, plus adjacent counties having a high degree of social and economic integration with the core as measured through commuting ties.	.D = Suppressed for confidentiality 1 = In a MSA 2 = Not in a MSA
MPC_YN	Metropolitan Principal City: An incorporated place or census designated place in a Metropolitan Statistical Area that meets specific population and workforce requirements.	.D = Suppressed for confidentiality 1 = In a Metropolitan Principal City 2 = Not in a Metropolitan Principal City

Additional geographies are identified through the intersection of CBSAFP_YN, METRO_YN, and MPC_YN shown in Table 18.

Table 18. Geographies Identified at the Intersections

Intersection	Additional Geography Level
CBSAFP_YN =1 and METRO_YN =2	Micropolitan Statistical Area: County or counties (or equivalent entities) associated with at least one urban cluster of at least 10,000 but less than 50,000 population, plus adjacent counties having a high degree of social and economic integration with the core as measured through commuting ties
METRO_YN =1 and MPC_YN=2	In an MSA, but not a Metropolitan Principal City: County or counties associated with at least one urbanized area of at least 50,000 population, plus adjacent counties having a high degree of social and economic integration with the core as measured through commuting ties, but is not incorporated place or census designated place within the Metropolitan Statistical Area.

Alternative and lower-level geographic identifiers are not included with the public use data file. Access to these variables is restricted to the Federal Statistical Research Data Centers (RDCs). Researchers can apply for RDC access; proposed projects must demonstrate scientific merit, require non-public data, be feasible, pose no risk to respondent confidentiality, and provide benefit to Census Bureau programs. The currently open RDCs are listed at <https://www.census.gov/about/adrm/fsrdc/locations.html>, and additional information on the RDC application process is available at <https://www.census.gov/about/adrm/ced/apply-for-access.html>.

Weighting Specifications

Overview

The NSCH uses child- and household-level weights for population-based estimates. These include

- Final Weight for Screened-in Households (FWH)
- Final Weight for Screener Children (FWS)
- Final Weight for Interviewed Children (FWC)

Each weight is the product of the base sampling weight, nonresponse adjustment factors, and raking adjustment (RAK). The FWC also includes a subsampling adjustment. Population controls are derived from applying demographic distributions from the 2019 American Community Survey (ACS) to the updated 2020 ACS population totals.

For 2021, the grouping process for the education dimension resulted in different state groupings from previous years. Additionally, one dimension, state by special health care needs (SHCN) status as determined from the screener, was removed from the final raking process (RAK). Oversampling in Stratum 1A (mostly children <5) reduced the proportion of SHCN among rostered children, thus the screener distribution of SHCN was no longer appropriate to include as a population control.

Base Weights (BW)

The weighting process began with the base sampling weight (BW) for each sample household. The base weight (i.e., sampling interval) for each sample housing unit was the inverse of its probability of selection for the screener. Base weights were calculated separately for each of the three strata and each state, including the District of Columbia. If there had been no nonresponse and the survey frame was complete, using this weight would give unbiased estimates for the survey population.

Screener Nonresponse Adjustment (SNA)

The Screener Nonresponse Adjustment (SNA) increases the weights of the households responding to the Screener to account for all the households not responding to the Screener.

Households were categorized into 24 groups to define the screener weight cells. Each cell was based on each combination of stratum (1A, 1B, or 2A), webgroup (High Web or High Paper), metropolitan statistical area status, and poverty indicator (the proportion of households with income less than 150% of the federal poverty level at the block group level).

SNA was calculated using the following formula:

$$\left(\frac{\text{summed BW of screener interviews} + \text{count of screener non-interviews}}{\text{summed BW of screener interviews}} \right)$$

where the count of screener non-interviews is an estimate of the expected number of eligible households (occupied, residential household) from those cases for which nothing is received back. The expected number of eligible cases is estimated by taking the eligibility rate among the known cases and applying it to the unknown cases.

The number of screener non-interviews was calculated using the following formula.

$$\left(\frac{\text{summed BW of screener interviews}}{\text{summed BW of screener interviews} + \text{summed BW of screener ineligible households}} \right) \times (\text{summed BW of households with unknown screener eligibility})$$

The resulting SNA was assigned to every household in the cell.

Household-Level Post-Stratification Adjustment Factor (HPSA)

All households with children that completed a screener were given a household-level weight. In addition to the base weight and screener nonresponse adjustment, a household post-stratification adjustment was applied in order to achieve the final household screener weight. This factor consisted of ratio adjustments to population controls at the household level obtained from the 2019 ACS data.

Households were put into one of 255 cells defined by state, race of the child selected for the topical, and Hispanic origin (yes or no) of the selected child if the selected child's race was White. Within each cell, the household post-stratification adjustment was calculated as the ACS population count for the cell divided by the cell's weighted total. The product of the base weight, screener nonresponse adjustment, and this household post-stratification adjustment constituted the final household screener weight.

First Raking to Population Controls: All Screener Children

All eligible children (four at most) from completed screener interviews were given a child-level screener weight. The weights of children from completed screener interviews were adjusted to match the 2019 ACS estimates for the following characteristics:

- Dimension #1 – State by Child's Race (White alone, Black alone, Asian alone, Other)
- Dimension #2 – State by Child's Ethnicity (Hispanic, Non-Hispanic)
- Dimension #3 – State by Child's Sex by Child's Age Group (0-5, 6-11, 12-17 years)

Each iteration of this process consisted of calculating three ratio adjustments, one for each dimension, sequentially. The adjustment factor calculated for Dimension 1 was applied to the weights accordingly and this newly adjusted weight went into the calculation of the adjustment factor for Dimension 2. This iterative raking process continued until the difference between the sum of the weights and the control total associated with each cell was less than 0.01% of the control. The resulting weight from this process was the final child-level screener weight for each eligible child. Only the children selected for the topical continued in the weighting process to eventually receive a final interviewed child weight.

Adjustment for Households with More than One Child

In households with multiple children, the selected child represented all eligible children in their household. Thus, a within-household subsampling factor was applied to account for the selection of a

single child, as well as the oversampling for young children and children with special health care needs (CSHCN). The value of this adjustment was the inverse of the probability of selection for the selected child. Probabilities varied by the number of children in the household, the presence of children aged 0-5, and the presence of CSHCN.

Adjustment for Topical Nonresponse

Similar to the screener nonresponse adjustment, the weights of the households responding to the topical needed to be increased to account for all of the households not responding to the topical. The adjustment considered all topical interviews (complete and sufficient partial) defined by questionnaires with valid answers for at least 25 of 50 test questions, and at least one item in Section H or beyond or the respondent submitted the topical portion of the web instrument. Returned topical that did not meet the criteria were considered a topical non-interview.

All topical-eligible households were put into one of sixteen cells depending on imputed poverty/non-poverty status, web group (High Web vs. High Paper), tenure (owner occupied or not), and presence of CSHCN. The topical nonresponse adjustment was calculated within each of the sixteen cells as:

$$\left(\frac{\text{weighted sum of topical interviews} + \text{weighted sum of topical non-interviews}}{\text{weighted sum of topical interviews}} \right)$$

After this adjustment, the selected children from topical non-interview households were no longer involved in the weighting process and only interviewed children continued to the last steps.

Raking Adjustment

The final step of the weighting process is accomplished through iterative raking to population controls attained from the ACS. The following eight analytical domains of interest were used:

- Dimension #1 – State by Household Poverty Ratio
- Dimension #2 – State by Household Size
- Dimension #3 – State Groupings by Respondent’s Education
- Dimension #4 – State by Selected Child’s Race
- Dimension #5 – State by Selected Child’s Ethnicity
- Dimension #6 – State by Selected Child’s Age Group
- Dimension #7 – National Selected Child’s Race by Ethnicity
- Dimension #8 – National Selected Child’s Sex by Single Age

The iterative raking process uses at most 100 iterations or until the weights converge to the population totals. Weights are converged when the absolute difference between the sum of the weights within each raking cell of all eight dimensions and the control total associated with each raking cell is less than one percent of the control.

For Dimension #3, some states needed to be grouped due to the low number of respondents in each state with less than a high school degree. States were grouped with others that had similar education distributions based on ACS data. The states were first sorted by the ACS-derived percent of children in

households where the respondent has less than a high school degree, followed by an additional sort by the percent of children in households where the respondent has a high school degree. State groupings were made with the intent of keeping these distributions similar within each group. The result was 11 state groupings and 26 stand-alone states. The following were the resulting groupings:

- Group 1: New Hampshire, North Dakota, and Vermont
- Group 2: Minnesota and Utah
- Group 3: Hawaii, Maine, and Montana
- Group 4: Connecticut, Massachusetts, and Virginia
- Group 5: South Dakota and Wyoming
- Group 6: DC and Maryland
- Group 7: Idaho and New Jersey
- Group 8: Delaware and North Carolina
- Group 9: Illinois and Washington
- Group 10: Missouri and Pennsylvania
- Group 11: Alaska and West Virginia
- Stand-alone states: Alabama, Arizona, Arkansas, California, Colorado, Florida, Georgia, Indiana, Iowa, Kansas, Kentucky, Louisiana, Michigan, Mississippi, Nebraska, Nevada, New Mexico, New York, Ohio, Oklahoma, Oregon, Rhode Island, South Carolina, Tennessee, Texas, and Wisconsin

Trimming of Large Weights

The resulting weights from each iteration of the final raking process were checked for extreme values to prevent a small number of cases with large weights from having undue influence on estimates and increasing the variance. An extreme value was determined to be a weight that exceeded the median weight plus six times the interquartile range (IQR) of the weights in each state. These extreme weights were truncated to this cutoff (median plus six times the IQR of weights in that state) and the weights were checked for convergence to the controls. Convergence required the weighted total of each cell to be within 1% of the control for the cell. If convergence was not met for every cell, another iteration of the raking process was applied again. This process of raking and trimming was reiterated until convergence was met and there were few extreme weights left. In general, the remaining extreme weights were observed to be very close to the cutoff. The remaining extreme weights were truncated a final time to the median plus six times the IQR in the state and the process was complete.

Population Controls

Population controls used throughout the weighting were derived from applying demographic distributions from the 2019 ACS one-year estimates to the 2020 ACS one-year estimates of population totals. Data collection for the 2020 ACS was interrupted by the Coronavirus pandemic. As such, the 2020 ACS one-year weights are considered experimental, but the population projections used to establish population totals are stable. Within this context, using the 2019 ACS population distributions on the 2020 ACS totals prioritizes methodological consistency while accounting for large-scale population change.

By using the ACS data, the weighted totals were ensured to match the population control totals available for key demographic variables for children and households in the U.S. The controls were used in the household post-stratification adjustment, the raking to attain the child-level screener weights, and the raking to attain the final topical interviewed children weights. Almost all controls used were at the state level, except for the last two dimensions in the second raking process, where national-level controls were used.

For the household post-stratification adjustment, the NSCH household weights were adjusted so that the sum of the weights equaled the ACS estimates for the number of households in each state by race (White, Black, Asian, Other) and by Hispanic origin (yes or no) if the selected child's race was White. In the first raking process, up to four children from each screener received adjustments so that the sum of the weights of all children listed on screeners equaled the ACS estimates for the number of children in each state by race, state by Hispanic origin, and state by sex by age group (0-5, 6-11, 12-17 years). Finally, in the second raking process, the weights of the NSCH topical interviewed children were adjusted so that the sum of their weights equaled the ACS estimates for each state by family poverty ratio ($\leq 100\%$, 101-200%, $>200\%$), household size (≤ 3 , 4, >4), respondent's highest level of education ($<$ High School, High School, $>$ High School), race, and Hispanic origin, as well as race by ethnicity and sex by age in years at the national level.

Limitations

To minimize the variability of the weights caused by large adjustment factors, cells having fewer than 30 cases were collapsed with a neighboring cell. The adjustment factors were then calculated for the merged cells by combining the population controls and the sample cases for the two cells. Since the individual cells were combined, and only one adjustment factor was created per cell, only the weighted total for the *combined* cell will match the control following the raking procedure. Consequently, the weighted totals for the individual cells will most likely not match the population controls for the original individual cells. As shown in Table 19, cells were collapsed in two of the dimensions in the last raking step.

Table 19. Collapsed Dimensions of Final Raking and Affected States

Collapse	Dimension Collapsed	Affected States
Black collapsed with Other in 14 states	Dimension #4 - State by Selected Child's Race (White, Black, Asian, Other)	AK, HI, ID, ME, MN, MT, NH, NM, ND, SD, UT, VT, WV, WY
Asian collapsed with Other in 20 states	Dimension #4 - State by Selected Child's Race (White, Black, Asian, Other)	AL, AR, DC, ID, IA, KS, KY, LA, ME, MS, MO, MT, NM, ND, SC, SD, UT, VT, WV, WY
Hispanic and Non-Hispanic collapsed in 2 states	Dimension #5 - State by Selected Child's Ethnicity (Hispanic, Non-Hispanic)	ND, VT

The NSCH public use file includes demographic detail that, in some cases, is partially adjusted in weighting. For any demographic characteristic not explicitly included in a post-stratification adjustment (i.e., the raking dimensions), the weights are not likely to sum exactly to that specific population total.

For example, in the national and state by race dimension of RAK, any child with American Indian/Alaska Native (AIAN) or Native Hawaiian/Other Pacific Islander (NHPI) as their race is included in the “Other” category. Thus, the sum of the AIAN or NHPI children’s weights will not usually sum to a population total of AIAN or NHPI but will sum in combination with all other races included in “Other” to the “Other” population control.

Along these lines, the strong state-level design of the NSCH means that median weights can differ significantly across states (i.e. larger weights for more populous states with a lower probability of selection given approximately equal state-level sample sizes). For any demographic characteristic not explicitly included in post-stratification adjustment by state, respondents for a minority population in a more populous state will have leverage on national estimates for that population that is not wholly captured in traditional measures of standard error. **We recommend caution when interpreting results for smaller population groups that are not included in post-stratification adjustment as they may not be state or nationally representative.** We advise researchers to consider alternative validations, including comparisons against unweighted estimates and nonparametric standard error estimation.

Estimation, Hypothesis Testing, and Data Use Guidelines

Variance Estimation

When survey weights are used, the resulting estimates from the 2021 NSCH are representative of all non-institutionalized children aged 0 to 17 years in the U.S. and in each state and the District of Columbia who live in housing units. These weighted estimates do not generalize to the population of parents, mothers, or pediatric health care providers. Analysts are advised to avoid statements such as “the percent of parents”.

Two stratum identifiers should be used to estimate variance: FIPSST (state of residence) and STRATUM (identifies households flagged with children). Each record in the data file is assigned a unique household identifier, HHID. Some analysts may be using statistical programs that only permit the specification of a single stratum variable. These users should define a new variable with 102 levels by crossing STRATUM (2 levels) with FIPSST (51 levels). This new variable can then be used as the stratum variable. For example, Stata users can specify only one variable in the strata() option of svyset. This new variable (named here as STRATACROSS) can be created using the following statement:

- `EGEN STRATACROSS = GROUP (FIPSST STRATUM)`

SUDAAN users can identify both FIPSST and STRATUM in the NEST statement. However, SUDAAN users should note that the first variable listed after the word NEST is assumed to be the stratum variable, and the second variable listed is assumed to be the PSU. To properly identify the PSU variable, the PSULEV option must be invoked in the NEST statement as shown here:

- `NEST FIPSST STRATUM HHID / PSULEV = 3;`

In both individual year and multi-year analyses, the NSCH sample size may be limited for smaller populations (e.g., American Indian or Alaska Native) and state-level subgroups or rare outcomes (e.g., adolescent CSHCN or autism in a particular state). Small sample sizes may produce unstable estimates. To minimize misinterpretation, we recommend only presenting statistics with a sample size or unweighted denominator of 30 or more. Further, if the 95% confidence interval width exceeds 20 percentage points or 1.2 times the estimate (\approx relative standard error $>30\%$), we recommend flagging for poor reliability and/or presenting a measure of statistical reliability (e.g., confidence intervals or statistical significance testing) to promote appropriate interpretation.

The Screener data file includes two weights. C_FWS is a child-level weight and is used for child-level estimates (e.g., the percent of children with special health care needs). FWH is a household-level weight and is used for household-level estimates (e.g., proportion of households with at least one child with special health care needs). Each row on the Screener data file represents a rostered child. When calculating household-level estimates, households should be collapsed to a single row (using HHIDS as the household ID) or should select a single row from each household (e.g., LINENUM = 1).

State-level estimates may be compared to national estimates using a nested z-test to identify significant differences at a given alpha or Type 1 error level (e.g., 0.05, 0.01). The formula for this is as follows:

$$Z = \frac{\bar{X}_i - \bar{X}_j}{\sqrt{SE_i^2 + SE_j^2 - 2P * SE_j^2}}$$

Where j is a subset of i (e.g., Alabama as part of the Total US), \bar{X} is the mean or proportion, SE is the standard error, and P is the proportion of the weighted denominator for a given indicator that is specific to j (e.g., Alabama weighted denominator divided by the Total US weighted denominator). A simple independent Z-test would be a more conservative test that may increase Type II error—the probability of failing to reject the null of no difference when there is a difference.

Statistical Testing for Overlapping Groups

The 2021 NSCH Topical public use data file includes three additional measures of the child’s reported race: SC_AIAN, SC_ASIAN, and SC_NHPI. The primary race variable that has been available on the NSCH public use data files since 2016, SC_RACE_R, identifies single race categories and collapses to a single “Two or More” category if the child is identified in two or more race groups. These new measures, on the other hand, indicate whether a child is American Indian/Alaska Native alone or in combination with another race (SC_AIAN), Asian alone or in combination with another race (SC_ASIAN), or Native Hawaiian/Other Pacific Islander alone or in combination with another race (SC_NHPI).

The addition of these ‘alone or in combination’ race indicators enables holistic examination for these population groups and substantially increases counts for indigenous populations, as more than half identify as multiple race. Ideally, group comparisons are made with mutually exclusive categories to avoid masking differences (e.g., AIAN alone or in combination versus another race group alone).

If comparisons are made between overlapping groups by race (e.g., comparing AIAN alone or in combination to NHPI alone or in combination, where an individual that identifies as AIAN and NHPI belongs to both groups), it should be noted that differences between the two groups will be smaller and standard statistical tests that assume independence will be more conservative than if the comparison were between mutually exclusive groups.

In certain cases, where one group is completely nested within a larger population (e.g. AIAN alone as a subgroup of AIAN alone or in combination), a standard error adjustment for nested comparisons is available (see above section on [Variance Estimation](#)) but is essentially equivalent to an independent group comparison (e.g. AIAN alone versus AIAN in combination)..

Combining Data across Survey Years

Data across multiple years of the redesigned NSCH (2016 and later) can be combined to increase the analytic sample size. By leveraging a larger sample, data users can analyze smaller population groups and rare outcomes that are not sufficiently represented in a single year sample and produce national and state-level estimates with smaller standard errors. Guidance for producing multi-year estimates is available at <https://www2.census.gov/programs-surveys/nsch/technical-documentation/methodology/NSCH-Guide-to-Multi-Year-Estimates.pdf>.

Confidentiality

Participation in the 2021 NSCH was voluntary, and all data collected that could potentially identify an individual person are confidential. Data are kept private in accordance with applicable law. Respondents are assured of the confidentiality of their replies in accordance with 13 U.S.C. Section 9. All access to Title 13 data from this survey is restricted to Census Bureau employees and those holding Census Bureau Special Sworn Status pursuant to 13 U.S.C. Section 23(c). In compliance with this law, all data released to the public are only in a statistical format. No information that could personally identify a respondent or household may be released. The Screener and Topical public use data files went through a thorough disclosure review process and were approved by the Census Disclosure Review Board prior to release.

Guidelines for Data Use

The U.S. Census Bureau is conducting the NSCH on the behalf of the Health Resources and Services Administration's Maternal and Child Health Bureau (HRSA MCHB) within the U.S. Department of Health and Human Services (HHS) under Title 13, United States Code, Section 8(b), which allows the Census Bureau to conduct surveys on behalf of other agencies. Title 42 U.S.C. Section 701(a)(2) allows HHS to collect information for the purpose of understanding the health and well-being of children in the U.S. The data collected under this agreement are confidential under 13 U.S.C. Section 9. All access to Title 13 data from this survey is restricted to Census Bureau employees and those holding Census Bureau Special Sworn Status pursuant to 13 U.S.C. Section 23(c).

Any effort to determine the identity of any reported case is prohibited. The Census Bureau and HRSA MCHB take extraordinary measures to assure that the identity of survey subjects cannot be disclosed. All direct identifiers, as well as 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
- Not link this data set with individually identifiable data from any other Census Bureau or non-Census Bureau data sets

Use of the data set signifies users' agreement to comply with the previously stated statutory-based requirements. Before releasing any statistics to the public, the Census Bureau reviews them to make

sure none of the information or characteristics could identify someone. For more information about the Census Bureau's privacy and confidentiality protections, contact the Policy Coordination Office toll-free at 1-800-923-8282.

Supporting Material

References

Bethell CD, Read D, Neff J, Blumberg SJ, Stein RE, Sharp V, Newacheck PW. 2002. "Comparison of the Children with Special Health Care Needs Screener to the Questionnaire for Identifying Children with Chronic Conditions—Revised." *Ambulatory Pediatrics*, Jan-Feb 2(1): 49-57.

Blumberg SJ, Luke JV. 2010. Wireless Substitution: Early Release of Estimates from the National Health Interview Survey, January–June 2010. National Center for Health Statistics. Available from: <http://www.cdc.gov/nchs/nhis.htm>

Blumberg SJ, Foster EB, Frasier AM, et al. 2012. Design and Operation of the National Survey of Children's Health, 2007. National Center for Health Statistics. *Vital Health Stat*, 1(55). Available from: http://www.cdc.gov/nchs/data/series/sr_01/sr01_055.pdf

Bramlett MD, Blumberg SJ, Ormson AE, et al. 2014. Design and Operation of the National Survey of Children with Special Health Care Needs, 2009–2010. National Center for Health Statistics. *Vital Health Stat*, 1(57). Available from: http://www.cdc.gov/nchs/data/series/sr_01/sr01_057.pdf

Brick JM, Williams D, Montaquila JM. 2011. "Address-Based Sampling for Subpopulation Surveys." *Public Opinion Quarterly*, 75(3): 409-28.

Dillman DA, Smyth JD, Christian LM. 2009. Internet, Mail and Mixed-Mode Surveys: The Tailored Design Method, 3rd edition. Hoboken, NJ: John Wiley & Sons.

Foster EB, Frasier AM, Morrison HM, O'Connor KS, Blumberg SJ. 2010. All Things Incentive: Exploring the Best Combination of Incentive Conditions. Paper presented at the American Association for Public Opinion Research annual conference, Chicago, IL.

Raghunathan TE, Lepkowski JM, Hoewyk JV, Solenberger PW. 2001. "A Multivariate Technique for Multiply Imputing Missing Values using a Sequence of Regression Models." *Survey Methodology*, 27: 85-95.

Raghunathan TE, Solenberger PW, Hoewyk JV. 2016. IVEware: Imputation and Variance Estimation Software User's Guide (Version 0.3). Ann Arbor, MI: Institute for Social Research, University of Michigan.

Schaefer JL, Graham JW. 2002. "Missing Data: Our View of State of the Art". *Psychological Methods*, 7(2): 147-77.

Attachment A: 2021 Estimated State-Level Sample Sizes

State sample sizes by stratum were determined using the following criteria. First, the Stratum 1 oversampling rates for each state were maximized such that the variance did not far exceed that of a design that sampled equally in the two strata. Second, the target number of topical interviews per state was adjusted until the total sample size was at the desired size. For 2021, approximately 300,000 addresses yielded 892 topical interviews per state at the median and 998 topical interviews per state on average.

Table A-1: Address Sample Size and Strata Distribution by State

State	Total Sample (approx.)	Stratum 1A	Stratum 1B	Stratum 2A
Alabama	6100	45.9%	27.8%	26.3%
Alaska	7200	41.6%	22.4%	36.0%
Arizona	5500	45.5%	29.8%	24.7%
Arkansas	6600	42.8%	27.2%	30.0%
California	5000	42.7%	30.3%	27.0%
Colorado	10500	28.9%	34.2%	36.9%
Connecticut	4200	43.9%	32.5%	23.6%
Delaware	5800	40.8%	28.9%	30.3%
DC	5300	39.5%	26.1%	34.4%
Florida	5700	43.8%	30.2%	26.0%
Georgia	10500	34.4%	32.6%	33.1%
Hawaii	8700	25.0%	10.7%	64.3%
Idaho	3800	46.9%	31.3%	21.8%
Illinois	4600	45.6%	30.6%	23.8%
Indiana	4700	47.7%	28.7%	23.6%
Iowa	4300	44.1%	29.8%	26.1%
Kansas	4600	49.1%	29.1%	21.8%
Kentucky	5300	41.5%	29.5%	29.0%
Louisiana	8700	37.5%	29.2%	33.4%
Maine	5200	44.5%	30.8%	24.7%
Maryland	4100	45.6%	32.4%	22.0%
Massachusetts	3600	42.2%	33.2%	24.6%
Michigan	4500	47.8%	30.9%	21.3%
Minnesota	3400	43.4%	31.2%	25.4%
Mississippi	7800	43.1%	25.3%	31.6%
Missouri	4500	44.7%	30.4%	24.9%
Montana	5300	39.7%	24.4%	35.9%
Nebraska	5800	40.3%	31.5%	28.2%
Nevada	6400	41.1%	25.3%	33.6%
New Hampshire	4200	40.8%	32.8%	26.4%

New Jersey	4500	41.6%	29.2%	29.2%
New Mexico	7600	39.1%	23.3%	37.6%
New York	5600	40.4%	26.6%	33.0%
North Carolina	5400	44.0%	30.5%	25.5%
North Dakota	4800	42.6%	25.6%	31.8%
Ohio	9700	34.7%	38.3%	26.9%
Oklahoma	6200	44.1%	26.0%	29.9%
Oregon	15500	28.9%	37.8%	33.3%
Pennsylvania	4300	44.9%	31.5%	23.6%
Rhode Island	5300	40.2%	30.2%	29.6%
South Carolina	5400	46.3%	30.1%	23.6%
South Dakota	4800	43.8%	26.4%	29.8%
Tennessee	5200	46.5%	30.8%	22.6%
Texas	6900	43.7%	26.4%	30.0%
Utah	4100	45.7%	24.8%	29.4%
Vermont	4100	40.6%	30.8%	28.6%
Virginia	4400	43.6%	30.1%	26.3%
Washington	4000	45.8%	30.7%	23.5%
West Virginia	6300	39.0%	26.2%	34.8%
Wisconsin	8400	33.3%	37.7%	29.1%
Wyoming	5800	40.8%	26.5%	32.7%

Attachment B: Probabilities for Selected Child

Respondents are given a household type (1, 2, 3A, 3B, 4, 5A, 5B, 5C, 6, 7, 8) based on the following variables from screener responses:

- TOTKIDS_E Total number of eligible children
- CHILDY0_5 Indicator child in 0-5 years old
- CHILDNO_5 Indicator child is not 0-5 years old
- TOTCSHCN Total number of special needs children
- TOTNON Total number of not special needs children

Each household type has rules of probability to select a child for the topical questionnaire. Some household types include an oversample for child selection based on age and special needs. Table B-1 shows each household type, their corresponding combination of variable, and a child's probability of selection from that household.

Table B-1: Household Type Assignment from the Values of Five Screener Variables

Household Type	Variables					Probability of Selection
	TOTKIDS_E	CHILDY0_5	CHILDNO_5	TOTCSHCN	TOTNON	
TYPE=1 → HHTYP_1	0 or blank	n/a	n/a	n/a	n/a	No Child
TYPE=2 → HHTYP_2	1	n/a	n/a	n/a	n/a	100% (Single Child)
TYPE=3A → HHTYP_3A	2	2	0	2	0	50%
		0	2	0	2	
				2	0	
				0	2	
TYPE=3B → HHTYP_3B	2	1	1	2	0	0-5 years old: 62% 6-17 years old: 38%
				0	2	
TYPE=4 → HHTYP_4	2	n/a	n/a	1	1	CSHCN: 64% non-CSHCN: 36%
TYPE=5A → HHTYP_5A	3	3	0	3	0	33%
		0	3	0	3	
				3	0	
				0	3	
TYPE=5B → HHTYP_5B	3	1	2	3	0	0-5 years old: 44% 6-17 years old: 28%
				0	3	
TYPE=5C → HHTYP_5C	3	2	1	3	0	0-5 years old: 38% 6-17 years old: 24%
				0	3	

TYPE=6 → HHTYP_6	3	n/a	n/a	1	2	CSHCN: 48% non-CSHCN: 26%
TYPE=7 → HHTYP_7	3	n/a	n/a	2	1	CSHCN: 39% non-CSHCN: 22%
TYPE=8 → HHTYP_8	≥4	n/a	n/a	n/a	n/a	25%

Attachment C: Weighted Response Rates by State

Table C-1. Weighted Response Rates by State

State	Resolution Rate	Screener Conversion Rate	Screener Completion Rate	Topical Conversion Rate	Topical Completion Rate	Interview Completion Rate	Overall Response Rate
United States	50.7%	98.9%	45.4%	80.4%	32.3%	79.5%	40.3%
Alabama	49.5%	99.2%	43.0%	74.9%	28.4%	74.3%	36.8%
Alaska	63.9%	99.4%	52.7%	76.4%	33.8%	75.9%	48.5%
Arizona	53.2%	99.0%	45.6%	78.9%	31.5%	78.1%	41.6%
Arkansas	53.2%	99.2%	45.8%	77.7%	30.5%	77.0%	41.0%
California	47.1%	98.7%	43.3%	75.9%	30.5%	74.9%	35.3%
Colorado	55.0%	99.1%	49.2%	81.7%	36.3%	80.9%	44.5%
Connecticut	52.2%	98.8%	47.8%	78.6%	35.9%	77.6%	40.5%
Delaware	48.6%	99.2%	44.5%	78.6%	31.3%	78.0%	37.9%
District of Columbia	51.3%	99.7%	47.4%	85.7%	39.1%	85.4%	43.8%
Florida	48.8%	99.0%	42.7%	77.8%	33.3%	77.0%	37.5%
Georgia	44.8%	98.8%	38.9%	76.7%	28.4%	75.7%	33.9%
Hawaii	62.4%	99.3%	57.1%	74.0%	42.1%	73.6%	45.9%
Idaho	59.4%	98.8%	53.2%	79.1%	38.3%	78.1%	46.4%
Illinois	51.1%	99.2%	46.6%	79.2%	35.1%	78.5%	40.1%
Indiana	53.3%	99.2%	47.9%	77.6%	32.7%	77.0%	41.0%
Iowa	62.5%	99.3%	59.1%	77.9%	36.9%	77.4%	48.3%
Kansas	53.5%	99.1%	48.2%	80.8%	36.1%	80.1%	42.8%
Kentucky	53.2%	99.1%	47.0%	78.9%	32.1%	78.3%	41.6%
Louisiana	45.2%	98.3%	36.2%	76.6%	26.7%	75.3%	34.0%
Maine	61.1%	99.4%	52.5%	80.1%	34.9%	79.6%	48.7%
Maryland	49.0%	98.9%	45.2%	74.8%	31.3%	74.0%	36.2%
Massachusetts	52.3%	98.9%	48.5%	77.4%	36.5%	76.5%	40.0%

Michigan	49.2%	98.8%	44.5%	79.0%	33.3%	78.1%	38.4%
Minnesota	60.3%	99.3%	56.5%	82.3%	40.1%	81.8%	49.3%
Mississippi	49.0%	99.3%	42.4%	79.3%	29.7%	78.7%	38.5%
Missouri	54.5%	98.8%	48.0%	80.4%	35.6%	79.4%	43.2%
Montana	65.8%	99.6%	59.7%	83.2%	41.4%	82.8%	54.5%
Nebraska	59.7%	99.3%	55.9%	82.2%	38.3%	81.6%	48.7%
Nevada	47.1%	98.6%	42.1%	73.4%	28.6%	72.4%	34.1%
New Hampshire	56.7%	99.4%	51.3%	82.3%	37.9%	81.8%	46.3%
New Jersey	48.5%	98.0%	45.0%	76.5%	34.2%	75.0%	36.4%
New Mexico	59.0%	99.6%	50.6%	76.7%	30.5%	76.4%	45.1%
New York	48.3%	98.7%	43.3%	75.5%	30.5%	74.5%	36.0%
North Carolina	48.7%	98.7%	42.0%	76.7%	29.4%	75.7%	36.8%
North Dakota	63.7%	99.5%	58.3%	77.4%	36.5%	77.0%	49.0%
Ohio	50.0%	99.1%	45.1%	77.8%	31.1%	77.1%	38.5%
Oklahoma	54.1%	99.4%	47.4%	77.4%	30.9%	76.9%	41.6%
Oregon	55.6%	99.1%	51.8%	81.8%	38.1%	81.1%	45.1%
Pennsylvania	53.2%	98.9%	48.3%	78.5%	34.1%	77.6%	41.3%
Rhode Island	46.7%	99.0%	43.5%	76.7%	29.6%	75.9%	35.5%
South Carolina	49.0%	99.3%	43.7%	78.4%	31.7%	77.8%	38.1%
South Dakota	63.9%	99.7%	59.4%	83.4%	38.2%	83.2%	53.2%
Tennessee	46.6%	99.0%	41.4%	76.6%	29.7%	75.8%	35.3%
Texas	44.1%	98.2%	37.3%	75.6%	26.1%	74.2%	32.7%
Utah	55.7%	98.6%	51.1%	80.6%	38.1%	79.5%	44.3%
Vermont	72.4%	99.7%	65.9%	82.5%	41.0%	82.2%	59.5%
Virginia	51.4%	98.9%	47.2%	79.2%	36.0%	78.3%	40.3%
Washington	55.4%	99.5%	51.8%	80.5%	38.4%	80.1%	44.4%
West Virginia	61.8%	99.5%	52.9%	77.7%	31.1%	77.3%	47.8%
Wisconsin	59.0%	99.3%	54.8%	82.6%	42.0%	82.1%	48.4%
Wyoming	62.9%	99.3%	51.6%	81.5%	32.7%	80.9%	50.9%