



U.S. Department of Commerce  
U.S. CENSUS BUREAU  
*census.gov*

October 7th, 2019

# 2018 National Survey of Children's Health

---

## Data Users Frequently Asked Questions (FAQs)

The U.S. Census Bureau reviewed this data product for unauthorized disclosure of confidential information and approved the disclosure avoidance practices applied to this release. CBDRB-FY19-ROSS-B0148

The questions and answers below summarize key information about the 2018 National Survey of Children’s Health (NSCH). A detailed report on the design and operation of the survey will be released online in Fall 2019.

Previously, the NSCH underwent a significant redesign. Further information on that redesign can be found in the 2016 NSCH Methodology Report (<https://www.census.gov/content/dam/Census/programs-surveys/nsch/tech-documentation/methodology/2016-NSCH-Methodology-Report.pdf>).

## Background

### Who sponsored the 2018 NSCH?

- Primary funder: United States Department of Health and Human Services (HHS), Health Resources and Services Administration’s (HRSA) Maternal and Child Health Bureau (MCHB)
- Additional funding for specific questions was provided by:
  - Centers for Disease Control and Prevention (CDC), National Center on Birth Defects and Developmental Disabilities (NCBDDD)
  - United States Department of Agriculture (USDA), Food and Nutrition Service
  - United States Environmental Protection Agency (EPA)

### Who conducted the 2018 NSCH?

- The United States Census Bureau, Associate Director for Demographic Programs on behalf of the United States Department of Health and Human Services (HHS), Health Resources and Services Administration’s (HRSA) Maternal and Child Health Bureau (MCHB)

### What is the goal of the 2018 NSCH?

- The NSCH is designed to produce national and state-level data on the physical and emotional health of American children 0 - 17 years old. Information is collected on factors related to the health and well-being of children, including access to and utilization of health care, receipt of care in a medical home, family interactions, parental health, school and after-school experiences, and neighborhood characteristics.

### How long has HRSA MCHB been collecting these data?

- Since 2003, the NSCH has provided both national and state-level prevalence estimates for a variety of child health and health care indicators. HRSA MCHB also fielded the National Survey of Children with Special Health Care Needs (NS-CSHCN) in 2001, 2005-2006, and 2009-2010. Its purpose was to assess the prevalence and impact of special health care needs among children and their families in the U.S.
- From 2016 onward, the purpose of the NSCH remains the same, but the two individual surveys were combined into a single survey conducted every year.
- Information about, and data from, the NSCH and NS-CSHCN prior to 2016 can be accessed through the **CDC’s National Center for Health Statistics** (NCHS - <http://www.cdc.gov/nchs/slits/nsch.htm>) or the **Data Resource Center for Child & Adolescent Health** (<http://www.childhealthdata.org/>).
- Information about, and data from, the 2016, 2017, and 2018 NSCH is available through **HRSA MCHB** (<https://mchb.hrsa.gov/data/national-surveys>) or the **Census Bureau** (<https://www.census.gov/nsch>).

## Why are data from the NSCH important?

- The NSCH is the only national and state-level survey on the health and well-being of children, their families, and their communities. The data collected are critical for providing states with unique information for the Title V Maternal and Child Health Services Block Grant program, state-level planning and program development, federal policy and program development, and general scientific research. These data are also used by families, child health advocates, policymakers, and others.

## What happened to the National Survey of Children with Special Health Care Needs (NS-CSHCN)?

- When HRSA MCHB undertook the redesign of the NSCH, they asked experts to consider the content of both the NSCH and the NS-CSHCN to support an annual fielding of a single survey.
- The revised NSCH was designed to maintain as many of the original items as possible from both surveys. The revised NSCH contains the Children with Special Health Care Needs (CSHCN) Screener (<https://www.ncbi.nlm.nih.gov/pubmed/11888437>) to support the continued availability of detailed, national and state-level estimates on the prevalence and impact of special health care needs among U.S. children and families. The **Data Resource Center for Child & Adolescent Health** (<http://childhealthdata.org/learn/NSCH>) provides information that can help users to compare content between the redesigned NSCH, previous NSCH, and previous NS-CSHCN, to understand which items have been maintained, altered, or dropped.
- The NS-CSHCN was last administered as an independent survey in 2009-2010.

## When will the new 2018 data be released?

- Data from the 2018 NSCH will be released in October 2019.
- Public Use Files for the 2016 NSCH and later are available through the **Census Bureau** (<https://census.gov/programs-surveys/nsch/data.html>). National-level micro-data files that include state-level identifiers will be released annually for the foreseeable future.
- The 2018 NSCH data release allows some state-level estimates; however, due to sample size limitations it may not be possible to release or produce state-level estimates for all variables and subgroups of children annually (see Estimation and Hypothesis Testing section). Users will be provided with instructions for generating multi-year estimates by combining data files (see Data Files section).

## SAMPLE DESIGN

### What was the 2018 sample design?

- A sample of 176,052 mailable household addresses was selected from the Census Master Address File (MAF).
- As in 2017, the Census Bureau's Center for Administrative Records Research and Applications (CARRA) developed an indicator based on multiple sources of administrative data, which was used to identify households with children to improve sampling efficiency. Addresses that were directly linked to a child under 18 years old (stratum 1) were oversampled. Among the remaining addresses, administrative records and small-area geographic characteristics were used to assign a probability of child presence. Addresses with a higher probability of child presence (stratum 2a) were sampled at a lower rate than stratum 1, while addresses with a very low probability of child presence (stratum 2b) were excluded from sampling. Stratum 2b represented no more than 5% of households with children in each state.
- State-level samples were allocated with the goal of producing an approximately equal number of completed questionnaires in each state and the District of Columbia.

## How was eligibility determined?

- A screener questionnaire was used to determine if the address represented an occupied residence and if there were eligible children ages 0-17 living at the sampled address. If responses to the initial questionnaire indicated that there were children in the household, then a more detailed, age-specific questionnaire was provided to the household.
- The screener instrument also included a series of health questions used to determine whether each eligible child could be classified as having a Special Health Care Need (<https://www.ncbi.nlm.nih.gov/pubmed/11888437>).

## How many children were selected per household as the subject of the questionnaire?

- Only one child per household was selected to be the subject of the detailed topical questionnaire in an effort to minimize respondent burden. Eligible children within households that completed a screener were subsampled to serve as the subject child for one of the three age-specific topical questionnaires: children aged 0-5, 6-11, or 12-17 years.
- Children with Special Health Care Needs (CSHCN) were oversampled (at 80%) in order to allow robust estimates of this critical population. A secondary oversample of children aged 0-5 years (at 60%) was applied because younger children are generally underrepresented in household surveys (as a result of being underrepresented in the administration records used to identify households with children) and to offset a potential age bias of the CSHCN oversample as older children are more likely to have special health care needs. More details on sampled child selection are available in the 2018 Methodology Report, *Subsampling Specifications: Selection of Sampled Child* (<https://www2.census.gov/programs-surveys/nsch/technical-documentation/methodology/2018-NSCH-Methodology-Report.pdf>).
- For households responding via the web instrument, the subsampling process occurred dynamically after the screener questions were completed. The instrument was programmed to subsample a single child and move into the age-specific questions immediately.
- For households responding via the paper instrument, the screener questionnaire was completed by the respondent and returned. A single child per household was similarly selected and the paper questionnaire with the appropriate age-specific questions was mailed back to the respondent for completion.

## QUESTIONNAIRE

### Where can I find copies of the questionnaires?

You can find copies of the 2018 household screener and age-specific topical questionnaires here: <https://mchb.hrsa.gov/data/national-surveys/questionnaires-datasets-supporting-documents> and <https://census.gov/programs-surveys/nsch/technical-documentation/questionnaires.2018.html>

### What kinds of questions are asked?

Survey topics include:

- Child and family demographics
- Physical and mental health status
- Health and functional status, including current conditions and related functional difficulties
- Health insurance status, type, and adequacy
- Access to and use of health care services

- Preventive medical and dental care, and specialty services received
- Medical home access
- Early childhood-specific information (0-5 years)
- Middle childhood and adolescent-specific information (6-17 years)
- Transition to adult healthcare services (12-17 years)
- Family health and activities
- Impact of child's health on family
- Parental health status
- Parent's perceptions of neighborhood characteristics
- Access to community-based services

## What were the major changes to the content of the NSCH since 2016?

Seven questions were added to the 2017 NSCH questionnaire. These questions were:

- COLOR ("Can this child identify the colors red, yellow, blue, and green by name?")
- MOLD ("DURING THE PAST 12 MONTHS, other than in a shower or bathtub, have you seen any mold, mildew or other signs of water damage on walls or other surfaces inside your home?")
- PESTICIDE ("DURING THE PAST 12 MONTHS, how often were pesticides used inside your residence to control for insects?")
- A1\_ACTIVE/A2\_ACTIVE ("Have you ever served on active duty in the U.S. Armed Forces, military Reserves, or the National Guard?")
- A1\_DEPLSTAT/A2\_DEPLSTAT ("Were you deployed at any time during (fill with SC\_NAME)'s life?")

Eighteen variables were added to the 2018 NSCH data file. These are:

- TENURE ("Is this house, apartment, or mobile home")
- BLOOD\_SCREEN ("Was this condition [blood disorder] identified through a blood test done shortly after birth?")
- CYSTFIB\_SCREEN ("Was this condition [Cystic Fibrosis] identified through a blood test done shortly after birth?")
- GENETIC\_SCREEN ("Was this [genetic] condition identified through a blood test done shortly after birth?")
- OVERWEIGHT ("Has a doctor or other health care provider ever told you that this child is overweight?")
- HOSPITALSTAY ("DURING THE PAST 12 MONTHS, was this child admitted to the hospital to stay for at least one night?")
- MEDHISTORY ("Did you and this child receive a summary of this child's medical history (for example, medical conditions, allergies, medications, immunizations)?")
- 11 items were added to assess early language development, including:
  - ONEWORD ("Is this child able to do the following: Say at least one word, such as "hi" or "dog"?")
  - TWOWORDS ("Is this child able to do the following: Use 2 words together, such as "car go"?")
  - THREEWORDS ("Is this child able to do the following: Use 3 words together in a sentence, such as "Mommy come now.?"")
  - ASKQUESTION ("Is this child able to do the following: Ask questions like "who," "what" "when" "where"?")
  - ASKQUESTION2 ("Is this child able to do the following: Ask questions like "why" and "how"?")
  - TELLSTORY ("Is this child able to do the following: Tell a story with a beginning, middle, and end?")
  - UNDERSTAND ("Is this child able to do the following: Understand the meaning of the word "no"?")
  - DIRECTIONS ("Is this child able to do the following: Follow a verbal direction without hand gestures, such as "Wash your hands.?"")
  - POINT ("Is this child able to do the following: Point to things in a book when asked?")

- DIRECTIONS2 ("Is this child able to do the following: Follow 2-step directions, such as "Get your shoes and put them in the basket.?"")
- UNDERSTAND2 ("Is this child able to do the following: Understand words such as "in," "on," and "under"?"")

Additional differences between the 2016, 2017, and 2018 NSCH data files are noted in the data crosswalk:

<https://census.gov/programs-surveys/nsch/technical-documentation/codebooks.html>

## DATA COLLECTION

### When were the data collected?

- Data were collected from June 2018 through January 2019.

### How often are data collected?

- Starting in 2016, the NSCH is fielded annually. The previous iterations of this survey were conducted every 4 years. To meet the needs of data users, a redesign of the survey was undertaken in order to support the release of new national- level estimates annually, and state-level estimates every 2-3 years.
- In 2016, the NSCH utilized an increased sample size in order to support state-level analyses with a single year of data collection. Starting in 2017, new national estimates are available each year and state-level estimates generally require 2-3 years of combined data (see Data Files and Estimation and Hypothesis Testing sections).

### How were the data collected?

- Households were presented with two modes for completing the 2018 NSCH:
  - A web survey that combined the screener and detailed topical questionnaires in a single instrument. After the respondent completed the screener portion, and if the household had at least one eligible child, the respondent was directed into one of the three age-based topical questionnaires, which asked detailed questions about one randomly selected child in the household. If the respondent preferred, an interview could also be completed over the phone by a telephone questionnaire assistance (TQA) agent using the web instrument.
  - A mailout/mailback self-administered paper-and-pencil interviewing (PAPI) screener instrument was provided in each mailing for 30% of the sample most likely to respond by paper<sup>1</sup>. The remaining addresses in the sample received a PAPI screener instrument beginning in the second non-response follow-up mailing. If the household screened into the study (had at least one eligible child) using the PAPI screener instrument, a separate PAPI detailed, age-specific topical questionnaire was mailed to the household for one randomly selected child.
- Invitation letters, toll-free TQA numbers, cash incentives, follow-up mailings, a certified mailing (without return receipt), and Spanish-language questionnaires were used to help maximize response rates.
  - Half of sampled addresses were randomly assigned to receive their initial invitation in an envelope with a certified mail sticker. The class of certified mail did not require that the recipient sign upon delivery.

---

<sup>1</sup>The 2018 NSCH Methodology Report, available in Fall 2019 at <https://census.gov/programs-surveys/nsch/technical-documentation/methodology.html>, offers additional details on the process whereby addresses are assigned to receive a paper screener in the initial mailing.

## Who was the respondent?

- A parent or other caregiver with knowledge of the health and health care of the sampled child in the household. For the 2018 NSCH detailed topical questionnaires, 63% of respondents were mothers (biological, step, foster, or adoptive), 28% were fathers (biological, step, foster, or adoptive), and 6% were other relatives or caregivers; the remaining respondents were not identified by the relationship to the child.

## What was the average survey completion time?

- Based on web keystroke data and paper cognitive testing results:
  - The estimated average survey length for households without children was less than 5 minutes.
  - The estimated average survey length for households with children (the screener and detailed topical questionnaire combined) was approximately 39 minutes.

## How many questionnaires were completed?

- Of the 176,052 sampled addresses in the 50 states and the District of Columbia, 40.2% were screened for age-eligible children. Of these households, 55.1% reported age-eligible children living or staying in the household. **38,140 screener questionnaires with children were completed.**
- From the sample of households with age-eligible children, **30,530 detailed topical questionnaires were completed.** The number of children with completed questionnaires per state ranged from 520 (District of Columbia) to 769 (Arkansas).
- Of the 30,530 completed detailed topical questionnaires, 77.9% were completed using the web instrument and 22.1% were completed using the paper instrument.

## What were the response rates?

Four different response rates were calculated based on the completion of the screener questionnaire, detailed topical questionnaire, and overall completion of the survey. The calculations for the Screener Completion Rate, Topical Completion Rate, Interview Completion Rate, and Overall Response Rate are detailed below.

- A completed screener must be received from a valid household at the address sampled, and 1) indicate that there are no children present or 2) provide a valid age for at least one child. Of the households that completed a screener questionnaire:
  - 53.9% are households with children that completed the screener instrument and are included on the screener data file.
  - 44.9% are households without children.
  - 1.2% are households with children that did not complete the screener instrument and are not included on the screener data file.
- Both complete and sufficient partial topical questionnaires are included on the detailed topical questionnaire data file. 30,530 households with children returned a complete or sufficient partial topical questionnaire.
  - A returned topical questionnaire is considered complete if at least 40 of 50 test questions have valid answers, and 1) the respondent completed at least one item in Section K (family income, household and family count) or 2) the web instrument was submitted by the respondent. 98.0% of returned topical questionnaires were complete.
  - A returned topical questionnaire is considered a sufficient partial if at least 25 of 50 test questions have valid answers, and 1) the respondent completed at least one item in Section H or beyond or 2) the topical portion of the web instrument was submitted by the respondent. 2.0% of returned

topical questionnaires were sufficient partials.

- 7,611 households with children completed the screener questionnaire but did not return a topical questionnaire or returned an insufficient partial topical questionnaire.
- For some addresses, no response was received so it was not possible to determine if the address represented an occupied residence. These addresses are classified as unresolved. Among these addresses, the proportion of addresses that are valid households was estimated using the household rate among the resolved addresses who did respond. The proportion of those households that include children was also estimated using the proportion of households that have children by state and stratum in the 2016 American Community Survey.
- The Screener Completion Rate is the estimated proportion of households (occupied residences) that completed a screener. The Topical Completion Rate is the estimated proportion of all households with children that completed a detailed topical questionnaire. The Interview Completion Rate is the probability that a household that initiates the survey will complete it. 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). The Overall Response Rate is the product of the Interview Completion Rate and the Resolution Rate (the proportion of addresses that are resolved – i.e., have sufficient information to determine if the address represents an occupied residence). In other words, it is the probability that an address progresses through the three major stages of survey completion – resolution, screener, and topical questionnaire.
- This approach yields a national weighted Screener Completion Rate of 49.8%, weighted Topical Completion Rate of 36.9%, **weighted Interview Completion Rate of 78.0%**, and **weighted Overall Response Rate of 43.1%**.

### Were incentives used to encourage participation?

- Yes. Survey methods research strongly supports the use of unconditional incentives to improve response rates in self-administered survey data collection. To improve the likelihood of response and reduce response bias in the 2018 NSCH, sampled addresses received a \$2 cash incentive, \$5 cash incentive, or they were part of the control group that did not receive a cash incentive (\$0).
- In the initial screener invitation, the household received one of the three incentive amounts. 45% of addresses received a \$2 incentive; 45% of addresses received a \$5 incentive; the remaining 10% of addresses represented the control group (no cash incentive).
- Addresses that returned a paper screener listing children in the household received up to four topical invitations. These invitations were spread across nine rounds of topical mailings.
- 90% of households received a \$5 incentive in the initial topical mailing. 10% of households were assigned to the control group (no incentive).
- An incentive experiment in the 2016 NSCH established the efficacy of unconditional cash incentives to encourage response. The control group in the 2018 NSCH was designed to allow continued monitoring of the effect of the incentive on sampled addresses.

### What was the potential impact of nonresponse bias on key survey estimates?

- Although incentive efforts did improve the response rate, nonresponse remained in the survey. Sampling weights were adjusted to account for this nonresponse. An extensive nonresponse bias analysis of the 2016 NSCH and 2017 NSCH is available (<https://census.gov/programs-surveys/nsch/technical-documentation/nonresponse-bias-analysis.html>) and indicated no systematic bias in the 2016 and 2017 estimates. The 2018 nonresponse bias analysis will be available in Fall 2019.

## In what languages were the questionnaires available?

- The NSCH paper and web instruments were both available in English and Spanish.

## DATA FILES

### Will it be possible to conduct trend analyses using data from the revised NSCH (2016, 2017, and 2018) and NSCH surveys conducted prior to 2016?

- No. Due to changes in the survey's mode of data collection and sampling frame, as well as adjustments to item wording where necessary, it will not be possible to compare estimates from the redesigned survey to those from previous iterations of the NSCH or NS-CSHCN or to conduct related trend analyses.
- The redesigned NSCH will support trend analyses beginning with data from 2016. Data from 2018 can be compared with data from 2016 and 2017, but not with prior years.

### Can I combine data from different years of the redesigned survey?

- Yes. The data sets can be combined (appended) to derive multi-year estimates, though the survey weights should be adjusted. Since each year is already weighted to represent the population of children residing in households for that year, the weight can simply be divided by the number of years being combined to derive multi-year estimates with an average annual or midpoint population size. For example, to calculate the combined 2016-2018 weight, analysts must divide each individual survey weight by 3 (i.e., number of survey years being combined; see <https://www2.census.gov/programs-surveys/nsch/technical-documentation/methodology/NSCH-Guide-to-Multi-Year-Estimates.pdf>).
- Note: When analyzing combined years of data, it is also recommended that 2017 and 2018 STRATUM=2a records be recoded as STRATUM=2 to ensure that the combined file is correctly treated as having two mutually exclusive sampling strata rather than three. Sampling strata need to be defined correctly in order to properly estimate design-adjusted variance and standard errors. Whereas the two state-level sampling strata in 2016 were STRATUM=1 and STRATUM=2, sampling in 2017 split Stratum 2 into Strata 2a and 2b, with no households selected from STRATUM=2b.

### How many complete questionnaires are included on the data files?

- A screener level data file includes records for 38,140 households, which includes all age-eligible children (0-17 years old) reported in each household. This totals 71,335 child records.
- Each of the 71,335 child records contains general demographic information and special health care needs status for the resident children in sampled households.
- One age-eligible child from each of the 38,140 households with children was selected for the detailed topical questionnaire. Of the eligible households, 30,530 households completed the detailed topical questionnaire for the selected child.
- A detailed topical questionnaire data file includes records for these 30,530 children. Each record contains all data for the child and the household in which the child resides, including the child's health, special health care need status, health care, family functioning, parental health, neighborhood and community characteristics, health insurance coverage, and demographics.

### Is it possible to link the screener and detailed topical questionnaire data files?

- To protect the privacy of respondents, it is not possible to link the screener file and detailed topical

questionnaire file. Relevant variables from the screener, such as child's age, sex, and special health care need status, are included on the topical questionnaire file. Further, there are variables on the topical questionnaire file that provide data users with information about the family and household.

### What analyses should utilize the screener file?

- The screener file must be used to estimate the proportion of households with Children with Special Health Care Needs (CSHCN) at the state and national levels. The screener file can also be used to estimate the prevalence of CSHCN at the state and national levels; data users should indicate the source (screener or topical file) of the prevalence estimates.
- The screener file also includes basic demographics for each eligible child.
- When using the screener file to calculate estimates for child level variables like race of child (RACER) or special health care needs status of the child (C\_CSHCN), the child weight should be used (C\_FWS). When calculating estimates for household level variables like primary household language (HHLANGUAGE), the household weight should be used (FWH).

### Where can I find more information about derived variables and the data suppressed to protect confidentiality?

- The 2018 NSCH Codebook details the variables names, definitions, and values included on the screener and detailed topical questionnaire files. The codebook will be available in Fall 2019. The data files went through extensive disclosure review in order to protect the privacy of all respondents.
- To protect the confidentiality of individual respondents and children, responses for certain variables were collapsed or suppressed.
- More details will be available in the Methodology Report for the 2018 NSCH, which is expected to be released in Fall 2019.

## SAMPLING WEIGHTS

### How were the data weighted?

A child weight was produced for the detailed topical questionnaire file (FWC), and child (C\_FWS) and household (FWH) weights were produced for the screener file. The weighting scheme began with a base sampling weight, which was the inverse of probability of the selection of the address. Adjustments to the base sampling weight included the following steps:

- Adjustments for screener nonresponse
- Eligible children from the screener questionnaire cases were raked to population controls from the 2017 American Community Survey.
- A within-household subsampling factor was applied to screener questionnaire cases
- Adjustments for detailed topical questionnaire nonresponse
- A raking adjustment that matched sample demographic estimates to various demographic controls

Raking of screener questionnaire cases was performed using the following three analytical domains of interest:

- Dimension #1 – State by Child's Race
- Dimension #2 – State by Child's Ethnicity
- Dimension #3 – State by Child's Sex by Child's Age Group

Raking of detailed topical questionnaire cases was performed using the following eight analytical domains of interest:

- 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 SHCN Status
- Dimension #7 – National Selected Child’s Race by Ethnicity
- Dimension #8 – National Selected Child’s Sex by Age

### **What was the source of independent population control totals for raking?**

- 2017 American Community Survey 1-year estimates and the screener data.

### **What were the final sampling weight variables?**

- FWC (Final Weight for Surveyed Children, detailed topical questionnaire file)
  - This weight should only be used for records on the detailed topical questionnaire file
- C\_FWS (Final Weight for Screened Children, screener file)
  - This weight should only be used for child level questions on the screener file
- FWH (Final Weight for Screened Households, screener file)
  - This weight should only be used for household level questions on the screener file

## **ESTIMATION AND HYPOTHESIS TESTING**

### **Whom does this sample represent?**

- When survey weights are used, the resulting estimates 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”.

### **Are special programs needed to analyze these data?**

- Yes. For proper variance estimation, programs must be able to account for the complex sample design structure. Examples include SUDAAN, Stata, SPSS Complex Samples module, WesVar, the SAS SURVEY procedures, and the base program R with the separate SURVEY package.

### **What variables should be used to estimate variance?**

- Stratum identifiers: FIPSST (state of residence) and STRATUM (identifies households flagged with children)
- Primary sampling unit: HHID (unique household identifier, topical data file), HHIDS (unique household identifier, screener data file)
- 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 STATE (51 levels). This new variable can then be used as the single stratum variable.
  - 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 (FIPSS STRATUM)

- EXAMPLE: SUDAAN users can identify both FIPSS 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 primary sampling unit (PSU). To properly identify the PSU variable, the PSULEV option must be invoked in the NEST statement as shown here:

NEST FIPSS STRATUM HHID / PSULEV = 3

## Can the data be used for subgroup analysis?

The procedure of keeping only select records and list-wise deleting other records is called subsetting the data. Most software packages that analyze complex survey data will incorrectly compute standard errors for subsetted data, because subsetting the data can delete important design information needed for variance estimation. Analysts should not subset the data, with one exception: Subsetting the survey data to a particular state does not compromise the design structure. Analysts interested in examining specific population subgroups (such as children living in poverty) must use the appropriate options in their software package (e.g., SUBPOPN in SUDAAN).

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.

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),  $\bar{X}$  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.

## WORKING WITH MISSING DATA

### How are missing data identified on the data files?

- The SAS data files for the NSCH include special missing value codes for analysts who may wish to differentiate between different types of missing values.
  - (.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-specific topical questionnaire.
- (.D) Cannot Disclose— The value is suppressed in order to protect the privacy of the respondent/child.

## Are any missing data imputed?

The following demographic variables were imputed for the purpose of raking during weighting procedures. The child demographic variables were imputed using hot-deck imputation while Adult 1 education, household size, and poverty ratio were imputed using sequential regression imputation methods. In addition to the single imputation for weighting, the poverty ratio was also multiply imputed using the same sequential regression imputation methods. Guidance for the analysis of multiply imputed poverty data is available at <https://www2.census.gov/programs-surveys/nsch/technical-documentation/methodology/NSCH-Analysis-with-Imputed-Data-Guide.pdf>.

- Household Tenure: TENURE
- Child’s sex: C\_SEX
- Child’s race, detailed: C\_RACE\_R

The following variables are derived from C\_RACE\_R and therefore use the imputed values:

- Child’s race, recoded: C\_RACER
- Child’s race, recode, Asian included: C\_RACEASIA
- Child’s race, recode, American Indian/Alaska Native included: C\_RACEAIAN
- Child’s Hispanic origin: C\_HISPANIC\_R
- Selected child’s sex: SC\_SEX
- Selected child’s race, detailed: SC\_RACE\_R

The following variables are derived from SC\_RACE\_R and therefore use the imputed values:

- Selected child’s race, recoded: SC\_RACER
- Selected child’s race, recode, Asian included: SC\_RACEASIA
- Selected child’s race, recode, American Indian/Alaska Native included: C\_RACEAIAN
- Selected child’s Hispanic origin: SC\_HISPANIC\_R
- Adult 1 highest completed year of school: A1\_GRADE
- Household size: HHCOUNT
- Poverty ratio: FPL\_I1-FPL\_I6

The public use files also include flag variables in order to indicate which records contained missing values and were imputed. The missing rates for each variable are listed below:

- Flag for Household Tenure: TENURE\_IF (0.43%)
- Flag for child’s sex: C\_SEX\_IF (0.21%)
- All imputed child race variables share the same flag: C\_RACE\_R\_IF (0.65%)
- Flag for child’s Hispanic origin: C\_HISPANIC\_R\_IF (0.93%)
- Flag for selected child’s sex: SC\_SEX\_IF (0.14%)
- All imputed selected child race variables share the same flag: SC\_RACE\_R\_IF (0.32%)
- Flag for selected child’s Hispanic origin: SC\_HISPANIC\_R\_IF (0.58%)
- Flag for adult 1’s highest completed year of school: A1\_GRADE\_IF (2.00%)
- Flag for household size: HHSIZE\_IF (1.90%)
- Flag for poverty ratio: FPL\_IF (15.34%)

Finally, information for about 7% of second adults in the households (A2) were reported in error. Reported values in these cases were removed and replaced using hot deck imputation. Cases with imputed values are indicated using the A2 imputation flag (A2\_IF). More details on the nature of the reporting error, identifying cases with erroneous data, and the imputation methodology are available in the 2018 Methodology Report (<https://www2.census.gov/programs->

[surveys/nsch/technical-documentation/methodology/2018-NSCH-Methodology-Report.pdf](https://www.census.gov/programs-surveys/nsch/technical-documentation/methodology/2018-NSCH-Methodology-Report.pdf)).

## GUIDELINES FOR DATA USE

The United States Census Bureau conducts the NSCH on behalf of 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 files. 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 these data files for statistical reporting and analysis only;
- Make no use of the identity of any person discovered, inadvertently or otherwise;
- Not link these data files with individually identifiable data from any other Census Bureau or non-Census Bureau data sets.

Use of the data files signifies users' agreement to comply with the above-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.

## FURTHER INFORMATION

### Whom do I contact if I have questions about the 2018 NSCH after I read this document?

- For further questions about the survey design, operation, and analysis, please send an email to [childrenshealth@census.gov](mailto:childrenshealth@census.gov).

### What is the suggested citation for this document?

- The United States Census Bureau, Associate Director of Demographic Programs, National Survey of Children's Health. 2018 National Survey of Children's Health Frequently Asked Questions. September 2019. Available from: <https://www2.census.gov/programs-surveys/nsch/technical-documentation/methodology/2018-NSCH-FAQs.pdf>.