2017 National Survey of Children's Health

Source and Accuracy Statement

U.S. Census Bureau
10/01/2018

United States® Census Bureau
U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov
@uscensusbureau

1.0 INTRODUCTION

The National Survey of Children's Health (NSCH) is being conducted by the U.S. Census Bureau for the U.S. Department of Health and Human Services' (HHS) Health Resources and Services Administration's (HRSA) Maternal and Child Health Bureau (MCHB). It is designed to provide national and state-level information about the physical and emotional health and well-being of children under the age of 18 living in mailable residential housing units in the United States, their families and their communities, as well as information about the prevalence and impact of children with special health care needs.

This Source and Accuracy Statement (S&A) provides an overview for the following phases of the 2017 NSCH survey cycle. Hopefully it will enable an understanding of the creation of the data files, as well guidance on their use.

2.0 SAMPLE DESIGN

2.1 Creation of the Sample Frame

The population of interest for the 2017 NSCH is all children under the age of 18, residing in the United States on the date of the survey. Among many other key elements, the survey frame was designed to identify households with children and to provide information about household access to the Internet, which was critical for data collection.

The 2017 NSCH sample frame was developed from two sources: the Edited Master Address File Extract (EDMAFX) created by the Demographic Statistical Methods Division (DSMD) of the Census Bureau, and a file of administrative flags that was created by of the Bureau's Center for Administrative Records Research and Applications (CARRA).

2.1.1 *Use of the Edited Master Address File Extract*

The Census Bureau's Master Address File (MAF) is an accurate and up-to-date inventory of all known living quarters in the United States, Puerto Rico, and associated island areas. It supports

most of the censuses and surveys that the Census Bureau conducts, including the decennial census, the American Community Survey (ACS), and ongoing demographic surveys. The content of the MAF includes mailing and location addresses, unit type attributes, geographic codes for areas such as state, county, census tract, and census block for each living quarters, and source and history data.

The EDMAFX is created at least once every year, specifically for use by DSMD's ongoing demographic surveys. Of importance to the 2017 NSCH is the assignment of a housing unit validity flag (VALDF17), resulting from filtering rules and processes implemented by DSMD.

The January 2017 version of the EDMAFX was used in the NSCH sample frame creation and consisted of 3,142 county and county equivalent MAFs rolled up to 51 state-level MAFs, which include the District of Columbia. Only records having VALDF17=1 (valid housing unit) were kept, with the unique identification variable MAFID[1] to match to CARRA's file of Administrative Flags.

2.1.2 *Use of CARRA's File of Administrative Flags*

All MAFIDs in the January 2017 MAF-X[2] were appended with flags (e.g. poverty, internet access, child present) from data sources such as the Numident and the ACS. This national file was matched to the EDMAFX to produce the sample frame.

2.1.2.1 Processing Overview of CARRA's 2017 NSCH File of Administrative Flags

The frame for all households with children came from three data sources: the Numident, a list of Social Security Number applicants with data updated from various administrative records, and the CARRA kidlink file. See Figure 1 for an overview of the process.

The Numident is based on all individuals who have been assigned Social Security Numbers. Demographic data from the Numident is updated from federal tax data and various administrative records. There were 85,189,798 children in the December 2016 Numident who would be aged 0–17 years on June 1, 2017.

To identify and sample households containing children in the Numident, the children in the Numident had to be connected to the households in which they live. This was done with the CARRA kidlink file. The CARRA kidlink file is a prototype linkage between children and parents based on Census and administrative records. The file uses data from Census surveys and federal administrative records to link children Protected Identification Keys (PIKs[3]) to parent PIKs. It identifies the parents of children in the Numident. The source data for the CARRA kidlink file are: the Census Numident, the 2010 Census Unedited File, the IRS 1040 and 1099

[1] Since MAFID cannot be released, similar household ID variables were created and placed on the Screener (HHIDS) and Topical Files (HHID).
[2] CARRA used different extracts of the January 2017 MAF in their processing, specifically the MAF-X and the MAF-ARF.
[3] CARRA uses an anonymous identifier called a PIK to link individuals across datasets while protecting their personally identifiable information.

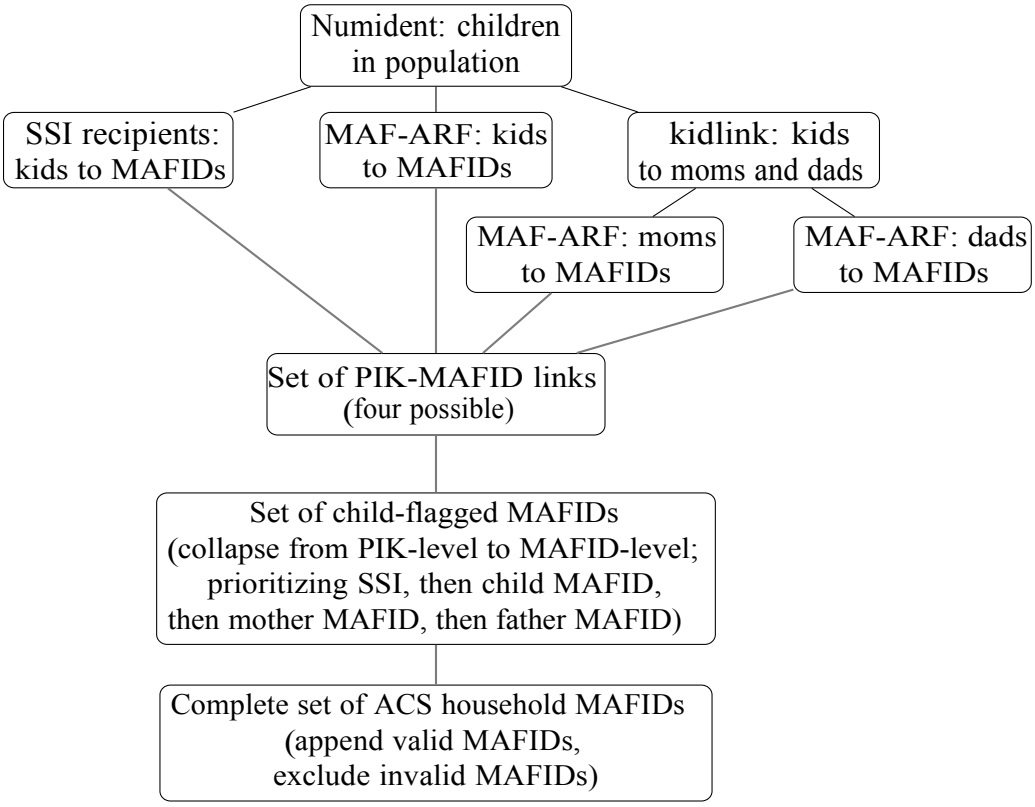
files, the Medicare Enrollment Database, the Indian Health Service Database, the Selective Service System, and Public and Indian Housing and Tenant Rental Assistance Certification System data from the Department of Housing and Urban Development. Of these, the IRS 1040 files provided the most significant information.

The MAF Auxiliary Reference File (MAF-ARF) was used to update household location. It links person identifiers to address identifiers using Census survey data and federal administrative data. The source data for the MAF-ARF file are the same as those listed for the CARRA kidlink file.

For each child observation from the Numident, there are four possible MAFIDs: the SSI MAFID, the kid to MAF-ARF MAFID, the child-to-kidlink-to-mother-to-MAF-ARF MAFID, and the child-to-kidlink-to-father-to-MAF-ARF MAFID. Using that order, a single MAFID was allocated. The MAFID match rate was 87.4 percent. The 74,464,032 children associated with a MAFID were then collapsed down to 38,716,792 unique MAFIDS. This implies 1.92 children per household for households assigned a flag.

The MAFID list was then scaled up to the universe of MAFIDs to allow sampling of unflagged households. A merge of the 38,716,792 unique child-flagged MAFIDS with the January 2017 ACS MAF-X file matched 38,716,730 MAFIDS with child flags and added 162,080,625 MAFIDs without child flags. Thus, the resultant file had 200,797,417 MAFIDS, of which 38,716,730 MAFIDs include child flags.

Figure 1: Illustration of CARRA's File Processing



2.1.2.2 Paper-only Response Probability Flag

Since 2012, ACS respondents have been able to submit survey forms over the Internet in addition to completing and mailing back a paper questionnaire. CARRA was able to use 2015 ACS response mode choices summarized at the block group level, as well as other block group and tract-level characteristics, to model Web and paper response mode probabilities by block group. Households on the NSCH frame were located within block groups and assigned a paper-only response probability.

Ultimately, a variable WEBGROUP was defined for the NSCH to distinguish households highly likely to complete the survey by paper (P) from those with a higher likelihood to complete by web (W). The 30% of households with the highest paper-only response probabilities were flagged as 'High Paper' (P) and received a paper questionnaire with the initial web invitation.

2.1.2.3 Local-area Household Income Relative to the Poverty Rate

The CARRA file also has a set of poverty variables from the 2015 5-year ACS file. These variables measure the proportion of households with household income in an interval defined by the poverty rate. Ultimately, a variable POVERTY was defined as Y or N from the proportion of households in the block group that have household income less than 150% of the poverty rate (30 percent cut-off).

2.1.3 *Final 2017 NSCH Sample Frame*

The data files detailed in Sections 2.1.1 and 2.1.2 were merged together based on MAFID to create the final sample frame.

2.2 Sampling Strata

Each state had three strata, 1, 2a, and 2b, which were defined by CARRA's child flag. Households flagged as having at least one child under the age of 18, determined by having an explicit link from a child to the household in administrative data, were assigned to Stratum 1. All other households which did not have explicit links to children were assigned to Stratum 2a or 2b based on their likelihood of having a child. Child presence in these households was modeled as a function of variables available in administrative data for all households on the MAF. The model was estimated with data from the most recent year of the ACS, in which child presence can be observed. Then, parameter estimates from that model were used to predict the likelihood of child presence for the households. These models were estimated separately for each state, and the threshold for bifurcation is based on an objective of maximizing the size of Stratum 2b while also maintaining 95% coverage of households with children in Strata 1 and 2a.

Variable state-level sampling occurred in only Strata 1 and 2a, with no households selected from Stratum 2b. Since Stratum 2b contains those households deemed very unlikely to have children, based on the lack of explicit links to children as well as the modeling results, the efficiency of the survey was increased by not sampling in the stratum.

2.3 Selection of the Sample Households and Additional Assignments

Recall that the 2017 NSCH sample frame is essentially the valid housing units from the MAF, appended with several administrative flags. Table 1 provides the calculated expected sample sizes, by state. Sample sizes were calculated to meet the goal of 433 Topical interviews per state, factoring in the expected valid address rate, expected United States Postal Service address standardization pass rates, response rates, and the prevalence of households with children. Addresses in Stratum 1 were sampled at a higher rate than Stratum 2a to increase the number of households with children in the sample while limiting the increase in the variance from the differential sampling rates. The oversampling factor (sampling rate for Stratum 1 divided by the sampling rate for Stratum 2a) ranged from 2 to 5 across the states. The total sample size was determined to be 170,728 housing units[4], 97,310 selected from Stratum 1 and 73,418 from Stratum 2a. (Note: The expected totals differ from the actual totals as a result of rounding in the sampling process. The resulting sample size was 170,726 households, with no state's actual sample size differing from its expected sample size by more than one.)

2.3.1 *Process of Selecting Households*

Sampling intervals determined the households selected to be in sample and were calculated for each of the two sampling strata in each state. The formula is the state-level stratum size on the frame divided by the calculated state-level expected sample size in the stratum.

When determining the random start for each stratum of each state, first the SAS function RANUNI(seed) was used to generate a number from the (0,1) uniform distribution. The returned value was then multiplied by the Sampling Interval to get the random start, or the first record to be in sample for that state and stratum.

2.3.2 *Assignment of Incentive and Infographic Group to the 170,728 Sample Records*

Incentive ($0 (control) or $2) and infographic group (receive infographic or not) for each MAFID were assigned randomly across the households that were selected for sample, by state. 90% of the sample was assigned to receive a $2 incentive, with the remaining 10% receiving no incentive and acting as a control to monitor the effectiveness of the incentive treatment. Within each of the two incentive groups, 50% of the sample received an infographic while the other 50% did not. These assignments for each of the sample records were made before any data was collected.

2.4 Selection of the Sample Children

2.4.1 *Determining Each Child's Eligibility*

A child is an eligible child if their age is less than 18 years.

[4] The total expected sample size of 170,728 was determined primarily from the available budget.

Table 1: 2017 NSCH Expected Sample Sizes, by Stratum and by State

State	Total Sample Size	Stratum 1 Sample	Stratum 2a Sample
Alabama	4,599	2,540	2,058
Alaska	5,757	1,785	3,972
Arizona	3,664	2,046	1,618
Arkansas	4,173	2,200	1,973
California	2,847	1,976	871
Colorado	2,680	1,633	1,047
Connecticut	2,769	1,759	1,011
Delaware	2,944	2,030	914
District of Columbia	3,226	2,091	1,135
Florida	3,621	2,454	1,167
Georgia	3,620	2,330	1,289
Hawaii	3,714	1,211	2,503
Idaho	2,840	1,578	1,262
Illinois	2,733	1,723	1,010
Indiana	3,188	1,975	1,213
Iowa	2,780	1,602	1,178
Kansas	2,877	1,907	970
Kentucky	3,336	2,044	1,292
Louisiana	4,787	2,836	1,951
Maine	3,726	1,805	1,921
Maryland	2,515	1,726	789
Massachusetts	2,474	1,573	901
Michigan	2,445	1,695	749
Minnesota	2,014	1,391	623
Mississippi	5,287	3,050	2,237
Missouri	3,037	1,967	1,069
Montana	3,765	1,698	2,067
Nebraska	2,601	1,657	944
Nevada	4,187	2,438	1,749
New Hampshire	3,119	1,770	1,349
New Jersey	2,664	1,782	881
New Mexico	5,379	2,096	3,283
New York	3,355	1,926	1,429
North Carolina	3,044	2,004	1,040
North Dakota	3,238	1,771	1,467
Ohio	2,712	1,883	829
Oklahoma	4,768	2,393	2,375
Oregon	2,367	1,573	794
Pennsylvania	2,496	1,691	806
Rhode Island	3,168	1,921	1,246
South Carolina	3,687	2,178	1,508
South Dakota	3,027	1,585	1,442
Tennessee	3,246	2,015	1,231
Texas	3,496	2,285	1,211
Utah	2,132	1,512	620
Vermont	4,470	1,547	2,923
Virginia	2,388	1,625	762
Washington	2,347	1,597	750
West Virginia	4,764	2,056	2,708
Wisconsin	2,104	1,489	615
Wyoming	4,555	1,889	2,666
National	**170,728**	**97,310 (57%)**	**73,418 (43%)**

2.4.2 *Determining the Status of each Eligible Child's Special Health Care Needs*

An eligible child in a household is deemed a child with special health care needs (C_CSHCN=1) if one or more of the following five groups have Screener responses of 'yes' to all of the questions in that group.

If:
Does (fill with CN_NAME) CURRENTLY need or use medicine prescribed by a doctor, other than vitamins? = yes (C_K2Q10=1) **AND**
Is (fill with CN_NAME)'s need for prescription medicine because of ANY medical, behavioral, or other health condition? = yes (C_K2Q11=1) **AND**
Is this a condition that has lasted or is expected to last 12 months or longer? = yes (C_K2Q12=1)

If:
Does (fill with CN_NAME) need or use more medical care, mental health, or educational services than is usual for most children of the same age? = yes (C_K2Q13=1) **AND**
Is (fill with CN_NAME)'s need for medical care, mental health, or educational services because of ANY medical, behavioral, or other health condition? = yes (C_K2Q14=1) **AND**
Is this a condition that has lasted or is expected to last 12 months or longer? = yes (C_K2Q15=1)

If:
Is (fill with CN_NAME) limited or prevented in any way in his or her ability to do the things most children of the same age can do? = yes (C_K2Q16=1) **AND**
Is (fill with CN_NAME)'s limitation in abilities because of ANY medical, behavioral, or other health condition? = yes (C_K2Q17=1) **AND**
Is this a condition that has lasted or is expected to last 12 months or longer? = yes (C_K2Q18=1)

If:
Does (fill with CN_NAME) need or get special therapy, such as physical, occupational, or speech therapy? = yes (C_K2Q19=1) **AND**
Is (fill with CN_NAME)'s need for special therapy because of ANY medical, behavioral, or other health condition? = yes (C_K2Q20=1) **AND**
Is this a condition that has lasted or is expected to last 12 months or longer? = yes (C_K2Q21=1)

If:
Does (fill with CN_NAME) have any kind of emotional, developmental, or behavioral problem for which he or she needs treatment or counseling? = yes (C_K2Q22=1) **AND**
Has his or her emotional, developmental, or behavioral problem lasted or is it expected to last 12 months or longer? = yes (C_K2Q23=1)

2.4.3 *Strategies for Selecting the 2017 NSCH Sample Children (SC_) from the Screener Responses*

For both the Paper and the Web data collection instruments, the sample child was selected from the first four eligible children after sorting by:

- special health care needs status
  - age (youngest to oldest)
- non-special health care needs status
  - age (youngest to oldest)

In the case of two or three children having the same age and the same special health care needs status, an additional sort by name (A to Z) was implemented. If they also had the same name, e.g., all 'blank', then sorting had no effect.

A special case was children in households that had four or more eligible children. These children were sorted first by their special health care needs status, then by name (A to Z), and then sorted by age (youngest to oldest).

A sample child was selected based on the criteria presented in Table 2. The strategies employed allowed for an oversample of both children with SHCNs and children aged 0 through 5 years.

Table 2: Strategies for Selecting the 2017 NSCH Sample Children (SC_)

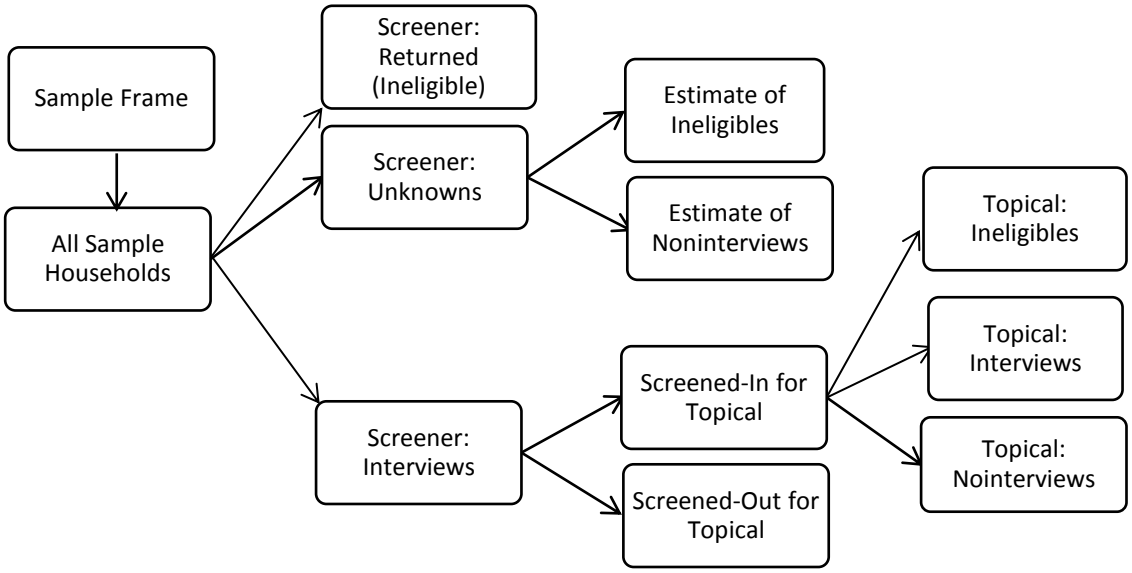
Number of Eligible Children in Household (TOTKIDS_R)	Number of Eligible Non-SHCN (TOTNONSHCN), CSHCN (TOTCSHCN)	% Probability of Selection for Non-SHCN	% Probability of Selection for CSHCN	Notes
1	1,0 or 0,1	100%		Single child is selected.
2	2,0 or 0,2	• If only 1 child is aged 0-5, that child's probability of selection is 62% and the other child's probability of selection is 38%. • Otherwise, each child has an equal chance of selection of 50%.		Includes 60% oversampling of children aged 0-5.
2	1,1	36%	64%	Includes 80% oversampling of CSHCN.
3	3,0 or 0,3	• If only 1 child is aged 0-5, that child's probability of selection is 44% and each of the other two children have an equal chance of selection of 28%. • If 2 children are aged 0-5, each has a probability of selection of 38% and the other child has a probability of selection of 24%. • If all 3 children are aged 0-5 or 6-17, then each child has an equal chance of selection of 33.3%.		Includes 60% oversampling of children aged 0-5.
3	2,1	52% • 26% probability of selection for *each* Non-SHCN child.	48%	Includes 80% oversampling of CSHCN.
3	1,2	22% • 39% probability of selection for *each* CSHCN.	78%	Includes 80% oversampling of CSHCN.
4 or more	Any combination	Before the sort, each of the first 4 children has an equal 25% probability of selection.		Simple random selection of 1 of the first 4 (sorted) children, regardless of Non-CSHCN or CSHCN.

3.0 SURVEY WEIGHTS

3.1 Overview of the Weighting Process

Figure 2 provides a framework for the weighting steps. The weighting process used the data from each phase of the data collection, from both the Paper and Web instruments, to produce final weights for the Screened-in Households, Screener Children, and Interviewed Children.

Figure 2: From Sample Frame to Final Outcome



The weighting process was done by state, with the District of Columbia treated as a state. Weighting for the interviewed children began with the base weight (BW) for each sample household, followed by a Screener nonresponse adjustment (SNA). Then, the eligible children from the Screener interview cases were raked to population controls (Child-Level Screener Factor=CLSF). A within-household subsampling factor (WHSF) was applied to the Screener interview cases, and a Topical nonresponse adjustment (TNA) was applied to the Topical interview cases. As a factor for the final weight for interviewed children, a final raking adjustment (RAK) to various demographic controls was performed. The weighting process for all Screener children was a subset of these six factors. Similarly, the screened-in households received a household-level weight, calculated using a small subset of the aforementioned factors as well as a Household Post-Stratification Adjustment (HPSA).

3.1.1 *Baseweight*

The BW for each sample housing unit is the inverse of its probability of selection for the Screener. Each state had two sampling strata with different probabilities of selection for each. If there had been no nonresponse and the survey frame was complete, using this weight would give unbiased estimates for the survey population.

3.1.2 *Screener Nonresponse Adjustment Factor*

The SNA increases the weights of the households responding to the Screener to account for all the households not responding to the Screener.

The count of Screener noninterviews is an estimate of the expected number of eligible households from those cases for which nothing is received. The term eligible here refers to the address belonging to an occupied, residential household. The expected number of eligible cases was estimated by taking the eligibility rate among the known cases and applying it to the unknown cases.

Sixteen Screener weighting cells were defined by the sampling stratum (STRATUM), a block-group poverty measure (yes/no) variable indicating the proportion of households with income less than 150 percent the poverty rate, an indicator of the likelihood of households to respond by paper (WEBGROUP), and a Metropolitan Area Flag (located within vs. outside of a metropolitan area).

Within each resultant Screener weighting cell, the SNA was defined as:

$$\left(\frac{\text{weighted sum of Screener interviews + S\_NONINT}}{\text{weighted sum of Screener interviews}}\right)$$

where S_NONINT =

$$\left(\frac{\text{weighted sum of Screener interviews}}{\text{weighted sum of Screener interviews + weighted sum of Screener ineligible households}}\right) \times$$

$$\text{(weighted sum of households with unknown Screener eligibility)}$$

This was the last of the weight processing for Screener households for which there was no Screener interview or interviewed households that indicated no eligible children.

3.1.3 *Household Post-Stratification Adjustment Factor*

All households who indicated on the Screener that there were eligible children present (also called screened-in households) were given a household-level weight. In addition to the BW and SNA, there was an HPSA applied in order to achieve the final screened-in household weight. This factor consisted of ratio adjustments to population controls attained from 2016 ACS data.

Households were put into one of 255 cells depending on their state, race of the selected child, and ethnicity of the selected child if the selected child's race was White. Cells were collapsed as necessary. Within each cell, the HPSA was calculated as the control for the cell divided by the cell's weighted total.

3.1.4 *Child-Level Screener Factor*

All eligible children (at most 4) from the Screener interviewed households were given a Child-Level Screener Weight in order to eventually produce state-level CSHCN prevalence estimates. This was accomplished through iterative raking to population controls attained from the ACS 2016 single-year estimates.

Raking to the population controls was accomplished using the following three analytical domains of interest, in this order: (Cells were collapsed as necessary.)

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

Each iteration consisted of three ratio adjustments. The ratio adjustments control the weights to the respective dimension control totals. Each ratio adjustment is called a rake. The first rake used the most recent intermediate weight (BW × SNA) as the child's input weight in the raking process. All subsequent rakes used the resulting weight from the previous rake as the input weight. The iterative raking process continued until convergence was met for all cells. Convergence required the cell's weighted total to be within 10% of the control.

At the end of the process, the CLSF was calculated as the weight after the final iteration divided by the weighted total prior to raking (BW× SNA).

Households where a child was selected from a completed Screener to receive a Topical interview, but become ineligible to complete a Topical were not assigned any further nonzero weighting factors. Examples may include households for which the Screener was received after the final Topical mailing, the child is no longer a resident of the household, etc.

3.1.5 *Within-Household Subsampling Factor*

Weights of the remaining eligible cases were adjusted for the subsampling of children within the households. The value of the adjustment is the inverse of the probability of selection for the selected children. Probabilities varied by the number of children in the household, the presence of children aged 0 through 5, and the presence of CSHCNs. The weights for the selected children now represented all children (at most 4) in the household, and took into account oversampling for CSHCNs and young children. See the details in the previous Table 2.

3.1.6 *Topical Nonresponse Adjustment Factor*

Similar to the SNA, the TNA increased the weights of the households responding to the Topical to account for all of the households not responding to the Topical. These households returned a Screener and went through the subsampling process to select a single child to be the subject of the Topical. If the respondent reached Section H and answered at least 50 percent of key items, then it was considered a Topical interview. A returned Topical that did not meet these conditions was considered a Topical non-interview.

Households were put into one of 8 cells depending on imputed poverty/non-poverty (yes/no), WEBGROUP (P/W), and presence of SHCN of the selected child. Within each of the 7 Topical weighting cells, collapsed as necessary:

$$\text{TNA} = \left(\frac{\text{weighted sum of Topical interviews} + \text{weighted sum of Topical Non-interviews}}{\text{weighted sum of Topical interviews}}\right)$$

Households for which there was no Topical interview were not assigned any further nonzero weighting factors.

3.1.7 *Raking Adjustment Factor*

This final step of the weighting process was accomplished through iterative raking to population controls attained from the ACS 2016 1-year estimates and the 2017 NSCH Screener data. Since the process was very similar to that of the CLSF, details are omitted in this section. The only significant differences were the addition of trimming and the dimensions:

- Dimension #1 – State by Household Poverty Ratio (≤1, (1,2], >2)
- Dimension #2 – State by Household Size (≤3, 4, >4)
- Dimension #3 – State Groupings by Respondent's Education (<HS, HS, >HS)
- Dimension #4 – State by Selected Child's Race (White, Black, Asian, Other)
- Dimension #5 – State by Selected Child's Ethnicity (Hispanic, Non-Hispanic)
- Dimension #6 – State by Selected Child's SHCN Status (Yes, No)
- Dimension #7 – Selected Child's Race by Ethnicity, at the National level (White Hispanic, White Non-Hispanic, Black Hispanic, Black Non-Hispanic, Asian, Other Hispanic, Other Non-Hispanic)
- Dimension #8 – Selected Child's Sex by Single Age, at the National level

3.1.8 *Trimming Extreme Weights*

At the end of each iteration, the weights were checked for extreme values. An extreme value was defined to be one that exceeded the median weight plus six times the interquartile range (IQR) of the weights in each state. These extreme weights were trimmed to this cutoff (six times the IQR of weights in that state). Then, the weights were checked for convergence, which required each cell's weighted total to be within 1% of the control for the cell. If convergence had not been achieved, the RAK raking steps were applied again and the new resulting weights were rechecked for extreme values and trimmed as before, continuing as was necessary until convergence was reached. At the end of the process, the RAK was calculated as the weight after the final iteration and trimming divided by the weighted total prior to raking (BW × SNA × SC_CLSF × WHSF × TNA).

Table 3 shows the distribution of the weights, by state, after the final iteration of raking and before the last and final trimming step. As shown by the low number of extremes in the final column, the proximity of the maximums to the cutoffs by state, and convergence to controls being met for all raking cells, it was decided to perform the final trimming at this point and the raking process was complete.

3.2 Final Weights Produced

Selected Child Weight (Topical) = FWC = BW × SNA × SC_CLSF × WHSF × TNA × RAK
Child Weight (Screener) = C_FWS = BW × SNA × C_CLSF
Household Weight (Screener) = FWH = BW × SNA × HPSA

3.3 Population Controls

The ACS is an ongoing national survey that samples approximately 3.5 million addresses annually, averaging about 290,000 addresses per month. These data are collected continuously throughout the year to produce annual population and housing estimates. The survey covers the resident population of the United States and Puerto Rico for people living in housing units and group quarters. (Note that the 2017 NSCH weighting cells only used the resident population of the United States for people living in housing units.)

The ACS produces critical information for small areas and small population groups – it is the only source of information for many of its topics in these small areas.

Two different sets of estimates, with weights, are released each Fall in the form of single-year (12 months of data) and 5-year (60 months of data) datasets. The 2017 NSCH weighting cells used the 2016 single-year ACS population controls.

Table 3: Summary of Last Raking Result before Final Trimming

STATE	MIN	Q1	MEDIAN	Q3	MAX	IQR	CUTOFF	Extremes
Alabama	475.0	1144.3	1790.4	2952.3	12675.3	1808.0	12638.6	3
Alaska	73.9	197.8	312.2	540.1	2375.8	342.3	2366.3	1
Arizona	648.7	1564.1	2471.7	4532.6	20369.9	2968.6	20283.0	1
Arkansas	296.8	739.4	1444.8	2556.1	12380.2	1816.7	12344.9	3
California	2059.4	7773.7	13340.2	24265.2	112246.7	16491.5	112289.1	0
Colorado	595.0	1356.2	1872.7	3426.5	14332.0	2070.3	14294.7	6
Connecticut	405.1	677.1	1069.0	1882.8	8403.8	1205.8	8303.6	4
Delaware	54.0	203.4	345.1	578.4	2576.0	375.0	2594.9	0
District of Columbia	19.2	86.7	143.5	313.9	1515.1	227.2	1506.8	2
Florida	974.2	3705.7	6125.4	12012.9	55666.0	8307.2	55968.3	0
Georgia	740.4	2475.5	4207.6	7061.7	31415.0	4586.2	31725.0	0
Hawaii	110.6	364.4	575.6	872.5	3623.9	508.1	3624.4	0
Idaho	224.3	528.2	887.4	1342.0	5773.7	813.8	5770.1	4
Illinois	1346.8	2992.3	4819.1	7637.5	32758.4	4645.3	32690.6	2
Indiana	585.4	1549.7	2516.5	4295.1	19047.3	2745.3	18988.5	5
Iowa	355.5	853.0	1266.4	2051.2	8490.8	1198.2	8455.6	4
Kansas	189.7	691.1	1059.4	1808.0	7889.5	1117.0	7761.3	16
Kentucky	583.4	1113.2	1801.2	3094.7	13711.5	1981.5	13690.4	3
Louisiana	398.5	1108.8	1681.3	3124.8	13800.0	2016.0	13777.1	4
Maine	126.4	263.8	398.8	655.2	2757.9	391.4	2747.1	4
Maryland	235.6	1170.0	2061.9	3574.1	16555.6	2404.1	16486.3	4
Massachusetts	684.8	1580.0	2369.1	3834.8	15978.5	2254.8	15898.1	10
Michigan	386.0	2125.0	3313.4	6206.6	27838.6	4081.6	27802.7	1
Minnesota	687.5	1287.0	2046.6	3784.6	17088.5	2497.6	17032.3	3
Mississippi	90.2	730.0	1214.7	2081.3	9188.2	1351.2	9322.1	0
Missouri	548.8	1386.2	2225.0	3736.8	16371.6	2350.7	16329.0	3
Montana	82.6	194.4	316.9	624.2	2852.1	429.9	2896.2	0
Nebraska	125.0	408.9	650.5	1425.8	6812.9	1016.9	6751.9	7
Nevada	189.7	599.5	988.2	1761.2	7985.8	1161.7	7958.5	4
New Hampshire	158.4	314.0	499.2	783.1	3108.7	469.1	3314.0	0
New Jersey	668.3	1686.6	2779.2	5179.8	24012.9	3493.3	23738.7	5
New Mexico	174.1	412.5	636.0	1299.2	5965.3	886.6	5955.7	18
New York	926.3	3901.2	6177.5	10443.6	45456.3	6542.3	45431.4	3
North Carolina	792.0	1966.7	3309.7	6638.7	31358.3	4672.0	31341.9	1
North Dakota	87.1	176.4	276.3	498.8	2211.7	322.4	2210.4	1
Ohio	1121.6	2699.5	4382.8	7287.5	32002.7	4588.1	31911.3	6
Oklahoma	313.5	995.0	1580.9	2719.0	11787.9	1724.1	11925.2	0
Oregon	431.1	907.9	1422.3	2720.4	12289.6	1812.6	12297.6	0
Pennsylvania	1373.7	2920.4	4335.8	6622.0	26610.2	3701.6	26545.4	9
Rhode Island	111.5	233.8	337.1	580.6	2423.7	346.8	2417.7	4
South Carolina	410.7	1253.9	1863.2	3119.8	13066.7	1865.9	13058.4	1
South Dakota	86.4	218.7	325.0	558.5	2373.0	339.8	2363.8	7
Tennessee	652.4	1726.0	2625.6	4458.6	19029.6	2732.5	19020.9	2
Texas	1199.2	7047.4	11351.7	21744.8	99817.3	14697.4	99536.1	4
Utah	394.4	981.2	1718.9	2649.7	11732.1	1668.5	11729.8	1
Vermont	30.9	127.4	203.3	326.2	1401.9	198.8	1396.1	2
Virginia	620.7	2008.9	3091.3	5115.3	21823.6	3106.4	21729.8	6
Washington	804.4	1740.1	2636.1	4585.1	19715.2	2845.0	19706.3	5
West Virginia	178.8	453.2	738.1	1245.7	5510.8	792.5	5492.9	2
Wisconsin	643.2	1521.6	2244.2	3729.5	15539.5	2207.9	15491.6	2
Wyoming	66.1	162.3	251.7	438.4	1917.3	276.1	1908.1	1

3.4 Checks to Guide the Use of the Three Final Weights

Using the assortment of Population Controls in the Attachment and the final weights in the NSCH files, the following are a few checks that the data user can do to more fully understand the use of the final weights:

- Check that the sum of the household weights for Screener interviews closely matches the control for each state (Column 2).
- Check that the sum of the Screener weights for children closely matches the control for each state (Column 3).
- Check that the sum of the Screener weights for female children closely matches the control for each state (Column 4).
- Check that the sum of the Topical weights for children with Poverty Ratio <=1 closely matches the control for each state (Column 5).
- Check that the sum of the Topical weights for children with SHCNs closely matches the control for each state (Column 6).

4.0 CALCULATING SAMPLING ERROR OF SURVEY ESTIMATES

4.1 Description of Sampling Error

The NSCH estimates are based on a sample; they may differ somewhat from the figures that would have been obtained if a complete census had been taken using the same questionnaire and instructions. This difference is known as sampling error and can be estimated from the survey data. While the simplest calculations of sampling error assume simple random sampling, these will underestimate the sampling error for the 2017 NSCH. This is because different sampling rates were used across the two sampling strata, as well as across states, resulting in a complex sample design.

Standard errors indicate the magnitude of the sampling error and can be used to construct confidence intervals around the survey estimates. By calculating the confidence intervals for a particular sample, one can say with a specified confidence that the average estimate derived from all possible samples is included in the confidence interval.

4.2 Estimating Sampling Error for the 2017 NSCH

Standard errors for the NSCH estimates can be obtained using the Taylor Series approximation method, which is available in software packages such as SAS, Stata, and SUDAAN. The sampling strata are identified by state and the child stratum flag, and the Primary Sampling Unit (PSU) is the household.

For SAS, the following statements are used:

- proc surveyfreq (or proc surveymeans or proc surveyreg)
- strata FIPSST and STRATUM
- cluster HHIDS (for the Screener) HHID (for the Topical)
- weight FWH (household weight), C_FWS (child weight), FWC (selected child weight)

For Stata the following statements are used:

- svyset strata FIPSST and STRATUM
- svyset psu HHIDS (for the Screener) or HHID (for the Topical)
- svyset pweight FWH (household weight), C_FWS (child weight), FWC (selected child weight)

For Stata, the two stratum variables need to be combined into a single variable.

For SUDAAN the following statements are used:

- proc …. design = WR;
- nest FIPSST STRATUM (HHIDS for the Screener or HHID for the Topical) / psulevel=3
- weight FWH (household weight), C_FWS (child weight), FWC (selected child weight)

For SUDAAN, the data file needs to be sorted by FIPSST and STRATUM, and then HHIDS (for the Screener) or HHID (for the Topical). HHID, HHIDS, FIPSST and STRATUM must be converted from character to numeric variable type.

5. Supporting Material

U.S. Census Bureau. Center for Administrative Records Research and Applications. "2017 National Survey of Children's Health Sample Frame." Unofficial document from Keith Finlay, dated April 13, 2017.

U.S. Census Bureau. "Sampling Specifications for the 2017 National Survey of Children's Health, including Creation of the Sample Frame." Forthcoming finalized memorandum from James B. Treat to Barry F. Sessamen.

U.S. Census Bureau. "Subsampling Specifications for the 2017 National Survey of Children's Health." Forthcoming finalized memorandum from James B. Treat to Barry F. Sessamen.

U.S. Census Bureau. "Weighting Specifications for the 2017 National Survey of Children's Health." Forthcoming finalized memorandum from James B. Treat to Barry F. Sessamen.

Attachment: Various Population Controls, by State

State	# of HHs with Children	# of Children	# of Female Children	# of Children in a HH with Poverty Ratio < 1	# of CSHCNs
Alabama	549,671	1,097,221	533,240	271,219	259,840
Alaska	85,742	186,806	87,638	26,789	33,390
Arizona	774,762	1,628,590	797,737	386,454	253,013
Arkansas	355,493	704,647	342,581	169,286	161,208
California	4,478,243	9,073,284	4,444,361	1,827,786	1,145,003
Colorado	652,979	1,257,916	609,752	171,436	230,495
Connecticut	393,462	750,763	366,958	97,525	160,268
Delaware	97,603	203,603	100,529	35,725	47,335
DC	60,059	120,168	59,729	30,648	21,111
Florida	2,037,785	4,133,530	2,022,894	877,005	830,010
Georgia	1,254,511	2,505,372	1,227,215	577,931	529,611
Hawaii	146,012	306,859	148,510	31,635	40,421
Idaho	192,321	435,239	211,282	77,600	79,204
Illinois	1,486,758	2,918,089	1,428,582	523,888	575,766
Indiana	781,867	1,571,675	766,614	312,194	328,677
Iowa	375,472	724,575	352,543	108,845	144,057
Kansas	340,960	712,627	346,389	101,638	148,567
Kentucky	530,631	1,009,088	492,946	253,311	236,736
Louisiana	540,797	1,115,188	543,928	320,376	252,112
Maine	130,762	254,167	124,082	44,054	58,174
Maryland	707,832	1,343,370	655,052	173,082	266,313
Massachusetts	753,987	1,372,524	671,796	187,248	276,612
Michigan	1,104,238	2,185,062	1,068,832	457,319	452,852
Minnesota	648,400	1,283,276	626,428	165,649	224,064
Mississippi	354,073	720,552	356,554	214,114	169,616
Missouri	711,997	1,384,804	673,313	266,476	280,359
Montana	111,255	227,405	108,812	34,774	45,323
Nebraska	235,352	472,493	231,413	67,548	86,058
Nevada	319,267	675,977	330,661	128,477	111,652
New Hampshire	144,686	259,127	124,253	20,605	54,629
New Jersey	1,042,981	1,979,553	968,833	292,985	320,536
New Mexico	226,067	487,713	239,142	147,656	89,382
New York	2,120,199	4,153,616	2,032,485	866,087	614,267
North Carolina	1,204,402	2,290,861	1,117,540	500,512	470,358
North Dakota	87,554	173,849	82,946	21,557	29,491
Ohio	1,357,635	2,604,031	1,275,576	536,182	547,170
Oklahoma	468,607	959,366	466,200	220,808	222,260

State	# of HHs with Children	# of Children	# of Female Children	# of Children in a HH with Poverty Ratio < 1	# of CSHCNs
Oregon	444,307	865,279	422,861	149,350	162,876
Pennsylvania	1,371,112	2,663,922	1,302,021	497,392	505,562
Rhode Island	111,843	207,806	101,752	34,827	42,616
South Carolina	547,272	1,096,675	536,852	254,432	226,443
South Dakota	99,601	212,742	102,938	35,460	35,058
Tennessee	769,190	1,500,270	728,168	343,850	289,084
Texas	3,521,198	7,275,916	3,563,040	1,648,612	1,074,686
Utah	391,243	918,765	446,446	105,230	143,037
Vermont	65,328	118,249	57,186	17,521	23,485
Virginia	989,439	1,863,541	911,000	272,282	391,380
Washington	834,283	1,626,757	795,247	225,070	305,432
West Virginia	199,883	375,537	185,645	91,471	89,603
Wisconsin	656,926	1,283,467	627,716	204,904	212,541
Wyoming	64,912	140,269	68,551	15,435	27,970
National	36,930,959	73,432,181	35,886,769	14,442,260	13,325,713