



# 2019 SURVEY DESCRIPTION

**Division of Health Interview Statistics  
National Center for Health Statistics  
Hyattsville, Maryland**

**Centers for Disease Control and Prevention  
U.S. Department of Health and Human Services**

**September 2020**

## Table of Contents

NCHS Website.....	4
NHIS Electronic Mail List.....	4
Questions about NHIS?.....	4
Guidelines for Citation of Data Source .....	5
Data User Agreement .....	6
Please Read Carefully Before Using the National Health Interview Survey .....	6
Terms and Conditions.....	6
Sanctions for Violating NCHS Data Use Agreement .....	6
What’s New in 2019? .....	7
About NHIS .....	8
NHIS Methods.....	9
I. Sample Design .....	9
II. Interviewing Procedures .....	10
III. NHIS Redesign .....	12
IV. Sponsors .....	17
V. Sample Sizes and Response Rates.....	18
VI. Weighting .....	20
VII. Variance Estimation.....	21
VIII. Editing the Data During and After the Interview.....	21
Orientation to How to Use NHIS data .....	24
I. Survey Data Files and Documentation .....	24
II. Using Weights and Variance Estimation Variables.....	30
III. Merging and Appending Files.....	36
IV. Section Acronyms and Questionnaire Order.....	45
Sample Adult’s Health .....	48
I. Health Status and Conditions .....	48
II. Functioning and Disability .....	50
III. Pain and Pain Management .....	53
IV. Health Care Access and Health Service Utilization.....	54
V. Preventive Care and Services .....	56
VI. Health-Related Behaviors.....	57
VII. Mental Health.....	58
Sample Child’s Health.....	60
I. Health status and conditions.....	60
II. Functioning and Disability .....	60
III. Health Care Access and Health Service Utilization.....	63
IV. Behavioral and Mental Health.....	64
V. Stressful Life Events.....	67

---

Health Insurance.....	68
Characteristics about the Sample Adult and Sample Child .....	72
Characteristics about the Parents of the Sample Child.....	76
Characteristics about the Spouse or Partner of the Sample Adult .....	78
Characteristics about the Family and Household of the Sample Adult and Sample Child.....	79
Telephone Usage .....	86
NCHS Data Linkage Program .....	87
Medical Expenditure Panel Survey (MEPS) .....	89
References.....	90
Appendix: Availability of Health, Health Insurance and Selected Demographic Questions Not Included in the Sample Adult and Sample Child Public Use Files.....	92

---

## NCHS Website

Data users can obtain the latest information about the National Health Interview Survey (NHIS) by periodically checking our website:

<https://www.cdc.gov/nchs/nhis.htm>

The website features downloadable data and documentation for the 2019 NHIS and previous years, as well as important information about any modifications or updates to the data and/or documentation. Published reports from previous years' surveys are also available, as are updates about future surveys and datasets. Data files and documentation can be found at:

<https://www.cdc.gov/nchs/nhis/data-questionnaires-documentation.htm>

## NHIS Electronic Mail List

Data users are encouraged to join the NHIS Listserv, an electronic mail list. The Listserv is made up of over 3,000 NHIS data users located around the world who receive e-news about NHIS surveys (e.g., new releases of data or modifications to existing data), publications, workshops, and conferences. To join, go to

[https://www.cdc.gov/nchs/products/nchs\\_listservs.htm](https://www.cdc.gov/nchs/products/nchs_listservs.htm), and select 'National Health Interview Survey' as one of your options, or click on "Contact Us" and then "Listserv" on the NHIS website.

## Questions about NHIS?

The staff of the Division of Health Interview Statistics at the National Center for Health Statistics respond to data users' questions about NHIS. Users may call us at 301-458-4901 and leave a voice message or e-mail us at [nhislist@cdc.gov](mailto:nhislist@cdc.gov). A response may take 1-2 business days.

---

## Guidelines for Citation of Data Source

With the goal of mutual benefit, the National Center for Health Statistics (NCHS) requests that recipients of NHIS data files cooperate in certain actions related to their use.

Any published material derived from the 2019 NHIS data should acknowledge “National Center for Health Statistics, National Health Interview Survey” as the original source. The full spelling of the source without the use of acronyms is preferred. The suggested citation to appear at the bottom of all tables and graphs is as follows:

- Data Source: National Center for Health Statistics, National Health Interview Survey, 2019

In a bibliography, the suggested citation for this document should read:

- National Center for Health Statistics. Survey Description, National Health Interview Survey, 2019. Hyattsville, Maryland. 2020.

The suggested citation for 2019 NHIS survey data and other documentation should read:

- National Center for Health Statistics. National Health Interview Survey, 2019. Public-use data file and documentation. <https://www.cdc.gov/nchs/nhis/data-questionnaires-documentation.htm>. 2020.

The published material should also include a disclaimer that credits the author’s analyses, interpretations, and conclusions to the author (recipient of the data file) and not to NCHS, which is responsible only for the initial data. Users who wish to publish a technical description of the data should make a reasonable effort to ensure that the description is consistent with that published by NCHS.

NHIS questionnaires are in the public domain and no permission is required to use them. Citation as to source, however, is appreciated.

Information on how to cite electronic media is available at: <https://www.cdc.gov/nchs/products/citations.htm>.

---

## Data User Agreement

### Please Read Carefully Before Using the National Health Interview Survey

The National Health Interview Survey (NHIS) is conducted by the National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention (CDC).

NCHS, CDC conducts statistical and epidemiological activities under the authority granted by the Public Health Service Act (42 U.S.C. § 242k). NCHS survey data such as NHIS are protected by Federal confidentiality laws including Section 308(d) Public Health Service Act [42 U.S.C. 242m(d)] and the Confidential Information Protection and Statistical Efficiency Act or CIPSEA [Pub. L. No. 115-435, 132 Stat. 5529 § 302]. These confidentiality laws state the data collected by NCHS may be used only for statistical reporting and analysis. Any effort to determine the identity of individuals and establishments violates the assurances of confidentiality provided by federal law.

### Terms and Conditions

NCHS does all it can to assure that the identity of individuals and establishments cannot be disclosed. All direct identifiers, as well as any characteristics that might lead to identification, are omitted from the dataset. Any intentional identification or disclosure of an individual or establishment violates the assurances of confidentiality given to the providers of the information. Therefore, users will:

1. Use the data in this dataset for statistical reporting and analysis only.
2. Make no attempt to learn the identity of any person or establishment included in these data.
3. Not link this dataset with individually identifiable data from other NCHS or non-NCHS datasets.
4. Not engage in any efforts to assess disclosure methodologies applied to protect individuals and establishments or any research on methods of re-identification of individuals and establishments.

By using these data, you signify your agreement to comply with the above-stated statutorily based requirements.

### Sanctions for Violating NCHS Data Use Agreement

Willfully disclosing any information that could identify a person or establishment in any manner to a person or agency not entitled to receive it, shall be guilty of a class E felony and imprisoned for not more than 5 years, or fined not more than \$250,000, or both.

---

## What's New in 2019?

- A redesigned NHIS questionnaire with new content and structure was implemented starting in January 2019. The interview begins with a household roster to ask basic demographic information for everyone in the household. One sample adult from each household is randomly selected to answer detailed questions about his or her health. One sample child, if present, is also randomly selected from each household and an adult knowledgeable and responsible for the child's health answers questions about the child's health. For additional information about the 2019 redesign, please visit: [https://www.cdc.gov/nchs/nhis/2019\\_quest\\_redesign.htm](https://www.cdc.gov/nchs/nhis/2019_quest_redesign.htm)
- Beginning with 2019, the NHIS annual data release will only include Sample Adult, Sample Child, Imputed Income – Sample Adult, Imputed Income – Sample Child, and Paradata files. The NHIS annual data release will no longer include household, family, or person data files. Household and family level information is present in the Sample Adult and Sample Child data files.
- There are four types of content included on the Sample Adult and Sample Child data files. Annual core is content that is included in the NHIS every year. Rotating core is scheduled to appear on a periodic basis of every other year, one out of every three years, or two out of every three years. Sponsored content includes questions that are funded by other agencies and are included when funding is provided. Sustaining sponsors fund content every year, whereas other sponsors fund content periodically. Finally, emerging content is new content that is sponsored by NCHS on emerging topics.
- Sponsored content in 2019 includes cancer control and prevention, immunizations, noncigarette tobacco product use, food security, food program participation, insulin use and arthritis.
- Emerging content in 2019 includes prescription opioid use and pain management.
- The NHIS includes some content that is asked at the family level. These questions may be answered by either the sample adult or sample child respondent if they are from the same family. If the sample adult and sample child are not from the same family, then both will receive the family-based questions.
- Two updates were made to the weighting process for the 2019 Sample Adult and Sample Child data files. These updates incorporate additional variables into the weighting process to better account for the decline in response to the NHIS over time. First, nonresponse weights are based on multilevel regression models that identify variables that predict the likelihood of responding to the survey and key health outcomes. Second, raking procedures are used to calibrate the nonresponse adjusted weights to population totals for age, sex, Hispanic or Latino origin and race, educational attainment, Census division, and Metropolitan Statistical Area (MSA) status.
- Imputed income variables are released concurrent with the Sample Adult and Sample Child data in 2019. The number of imputations was increased from five to 10 to allow more precise estimates. The Sample Adult and Sample Child Imputed Income files include 10 sets of continuous and categorical top-coded family income and poverty ratio variables.
- Demographic information of up to two parents residing with the sample child, and about the sample adult's spouse or cohabiting partner in the home are now included in the files. The urbanization level of where the sample adult and sample child live are also now included in the data files.

---

## About NHIS

NHIS is the principal source of information on the health of the civilian noninstitutionalized population of the United States and is one of the major data collection programs of the NCHS. The National Health Survey Act of 1956 provided for a continuing survey and special studies to secure accurate and current statistical information on the amount, distribution, and effects of illness and disability in the United States and the services rendered for or because of such conditions. The survey referred to in the Act, now called the National Health Interview Survey, was initiated in July 1957. Since 1960, the survey has been conducted by NCHS, which was formed when the National Health Survey and the National Vital Statistics Division were combined.

The main objective of the NHIS is to monitor the health of the United States population through the collection and analysis of data on a broad range of health topics. A major strength of this survey lies in the ability to categorize these health characteristics by many demographic and socioeconomic characteristics.

NHIS data are used widely throughout the Department of Health and Human Services (HHS) to monitor trends in illness and disability and to track progress toward achieving national health objectives. The data are also used by the public health research community for epidemiologic and policy analysis of such timely issues as characterizing those with various health problems, determining barriers to accessing and using appropriate health care, and evaluating Federal health programs.

Since 1957, the content of the survey has been updated about every 10–15 years to incorporate advances in survey methodology and coverage of health topics. In January 2019, NHIS launched a redesigned content and structure that differs from the 1997–2018 NHIS.



---

## NHIS Methods

### I. Sample Design

NHIS is a cross-sectional household interview survey. The target population for the NHIS is the civilian noninstitutionalized population residing within the 50 states and the District of Columbia at the time of the interview. The NHIS universe includes residents of households and noninstitutional group quarters (e.g., homeless shelters, rooming houses, and group homes). Persons residing temporarily in student dormitories or temporary housing are sampled within the households that they reside in permanently. Persons excluded from the universe are those with no fixed household address (e.g., homeless and/or transient persons not residing in shelters), active duty military personnel and civilians living on military bases, persons in long-term care institutions (e.g., nursing homes for the elderly, hospitals for the chronically ill or physically or intellectually disabled, and wards for abused or neglected children), persons in correctional facilities (e.g., prisons or jails, juvenile detention centers, and halfway houses), and U.S. nationals living in foreign countries. While active-duty Armed Forces personnel cannot be sampled for inclusion in the survey, any civilians residing with Armed Forces personnel in non-military housing are eligible to be sampled.

Because the NHIS is conducted in a face-to-face interview format, the costs of interviewing a large simple random sample of households and noninstitutional group quarters would be prohibitive; randomly sampled dwelling units would be too dispersed throughout the nation for cost-effective interviewing. To keep survey operations manageable, cost-effective, and timely, the NHIS uses geographically clustered sampling techniques to select the sample of dwelling units for the NHIS. The sample is designed in such a way that each month's sample is nationally representative. Data collection on the NHIS is continuous, i.e., from January to December each year.

The sampling plan is redesigned after every decennial census. A new sampling plan for the 2016–2025 NHIS was designed with results of the 2010 decennial census. The sampling process starts with partitioning the United States into 1,689 geographic areas. These geographic areas are defined as counties, county equivalents, or groups of counties, are almost always contiguous, and do not cross state boundaries. Next, within some states, the geographic areas are divided into two strata defined by population density (generally, urban counties and rural counties). For the remaining states, all the geographic areas form one stratum. Clusters of addresses were then defined within each stratum. The sizes of the clusters correspond generally to the size of an interviewer's workload over the course of the 10-year sample design period; the approximate size is 2,500 addresses per cluster. Each cluster is located entirely within one of the 1,689 originally defined geographic areas. Within each stratum, a specific number of clusters is systematically selected for the NHIS sample. The number selected is generally proportional to the number of clusters in the strata, e.g., larger strata have more clusters selected within the strata. The exception is in the 10 least populous states and the District of Columbia, where a slightly higher number of clusters are selected in order to ensure that all states have a minimum number of addresses in the sample.

Commercial address lists were used as the main source of addresses, supplemented by field listing. As of the beginning of 2018, the NHIS sampling frame consists of two non-overlapping parts: the unit frame (a list of addresses purchased from a vendor), and the area frame (generated by traditional field enumeration). Approximately 11% of the counties in the sample were part of the area frame. These area frame counties are typically counties with relatively few city-style addresses, and counties where the unit frame did not have acceptable coverage, i.e., where the vendor-supplied list did not adequately include all eligible households. The college dormitory sampling frame implemented in 2016–2017 was discontinued in 2018, and residents of

dormitories were eligible to be sampled in the non-college households where they reside via the household rostering.

When the sample design was redesigned for the 2016–2025 NHIS, the base of approximately 58,800 addresses was expected to yield approximately 27,000 adult and 9,000 child completed interviews in 35,000 households. Given the questionnaire redesign in 2019 and the 2019 response rates, the base sample is now expected to yield approximately 28,800 adult and 8,400 child completed interviews in 30,000 households annually.

The final sample in 2019 consists of 534 clusters of addresses; these clusters are found in 310 geographic areas, with clusters selected in every state. Note that, unlike previous NHIS sample designs, the geographic areas were not sampled directly in the 2016 design. Rather, the locations of the selected address clusters determined which geographic areas were included in the sample.

A report with further information on the 2016–2025 sample design is forthcoming.

## II. Interviewing Procedures

The U.S. Census Bureau, under a contractual agreement, is the data collection agent for the National Health Interview Survey. NHIS data are collected continuously throughout the year by Census interviewers. Nationally, about 750 interviewers (also called “Field Representatives” or “FRs”) are trained and directed by health survey supervisors in the U.S. Census Bureau Regional Offices to conduct interviews for NHIS. Interviewers are observed by supervisors periodically and their work is monitored by the Census Bureau’s PANDA system, a performance and data analysis program that provides monthly checks on response rates, completion rates, item response times, item nonresponse, telephone usage rates, and other data quality indicators. The supervisors responsible for the NHIS are career Civil Service employees who are selected through an examination and testing process. Interviewers receive thorough refresher training annually and other training during the year in basic interviewing procedures and in the concepts and procedures unique to the NHIS.

Each household address selected for participation in the NHIS is mailed a letter prior to the interviewer’s visit. The “advance letter” is mailed one week prior to the start of the interview period (one week before the 1st of the month) with the goal that it might be fresh on people’s mind when the FR makes contact the first few days of the month. This “advance letter” contains information about the purpose of the NHIS and the amount of time the interview will require, and it assures potential respondents that participation in the NHIS is voluntary. It also informs respondents that the information they provide is protected by law and details how the information will be used. When the interviewer arrives at the household address, he/she provides another copy of the “advance letter” to each respondent and obtains verbal consent for survey participation. A copy of the current “advance letter” and other NHIS materials available for distribution by FRs in the field are available at the NHIS participants page: <https://www.cdc.gov/nchs/nhis/participant.htm>.

The NHIS is conducted using computer-assisted personal interviewing (CAPI). The CAPI data collection method employs Blaise computer software that presents questions on computer screens to each interviewer. The instrument guides the interviewer through the questionnaire, automatically routing the interviewer to appropriate questions based on answers to previous questions. Interviewers enter survey responses directly into the computer, and the CAPI program determines if the selected response is within an allowable range, checks it for consistency against some of the other data collected during the interview, and saves the responses into a survey data file. The computer contains help facilities to aid interviewers in administering the CAPI

questionnaire. This data collection technology reduces the time required for transferring, processing, and releasing data, and it ensures the accurate flow of the questionnaire.

Face-to-face interviews are conducted in respondents' homes, but follow-ups to complete interviews may be conducted over the telephone. A telephone interview may also be conducted when the respondent requests a telephone interview or when road conditions or travel distances would make it difficult to schedule a personal visit before the required completion date. In 2019, 34.3% of the Sample Adult interviews and 31.7% of the Sample Child interviews were conducted at least partially by telephone.

For the Household Roster section of the questionnaire, any responsible household member aged 18 years or over is identified to act as the "household respondent." The household respondent provides names, age, sex, race, and ethnicity for all household members. The highest level of education completed and active military status is asked for all adult household members age 18 years or over. In addition to collecting this basic demographic information, the household roster interview also identifies whether all persons in the household are members of the same or different family. Note that in a multi-family household, a single "household respondent" provides household information for all families.

NHIS has consistently defined a family as an individual or a group of two or more people residing together who are related by birth, marriage, or adoption. A family additionally includes any unrelated children who are cared for by the family (such as foster children) and any unmarried cohabiting partners and their children. After the household roster is completed, data are collected on one adult and child per household. A "sample adult" is randomly selected by the computer from each household with at least one household member aged 18 years or over and is asked more detailed health related questions. The sample adult responds for him/herself to the questions in that section unless he/she is physically or mentally unable to do so, in which case a knowledgeable proxy may answer for the sample adult. Students aged 18 and over living away at college, trade, or commercial schools in on-campus housing are eligible to be interviewed in the location they consider to be their usual residence, such as their parent's or other family member's household. Students living away at school or college in off-campus housing will not be included as members of the household, since they could be sampled at their off-campus location. A "sample child" is randomly selected by the computer from each household with at least one child 17 years of age or younger. An adult respondent who was previously indicated to be knowledgeable and responsible for the sample child's health will be asked questions about that child. In 2019, 93.4% of the sample child respondents were the child's parent, either a biological, adoptive or stepparent. For each sampled household address, interviewers also maintain electronic documentation about the NHIS interview process, including contact attempts, observed characteristics about the exterior of the sample unit or vicinity, and descriptive information about the interview outcome.

All information collected by the NHIS that would permit identification of the individual is held strictly confidential, seen only by persons who work on the NHIS (including related studies carried out by the Public Health Service) with a need to know, and such information is not disclosed or released to anyone for any other purpose without the consent of the respondent. NCHS must adhere to Section 308(d) of the Public Health Service Act (42 U.S.C. 242m(d)), which forbids the disclosure of any information that may compromise the confidentiality promised to survey respondents. In addition, confidentiality protections are also mandated by the Confidential Information Protection and Statistical Efficiency Act of 2018 (Title III, Public Law No. 115-435).

Further information about data collection procedures is available in the 2019 NHIS Field Representative's (FR) Manual available on the NHIS website, <https://www.cdc.gov/nchs/nhis.htm>.

### III. NHIS Redesign

In 2019, the structure and content of the NHIS were redesigned to better meet the needs of data users, the Centers for Disease Control and Prevention, and the Department of Health and Human Services (DHHS). The goals of the redesign were to reduce respondent burden by shortening the length of the questionnaire, harmonize overlapping content with other federal health surveys, establish a long-term structure of ongoing and periodic topics, and incorporate advances in survey methodology and measurement.

The public was involved in the redesign process through public comments received through separate NCHS requests for input in 2015, 2016, and 2017. Additionally, technical expert panels consisting of subject matter experts in the fields of child health, chronic pain, injury, and income were convened to offer information about the directions and needs of each health-related field. For additional information about the 2019 redesign, visit: [https://www.cdc.gov/nchs/nhis/2019\\_quest\\_redesign.htm](https://www.cdc.gov/nchs/nhis/2019_quest_redesign.htm)

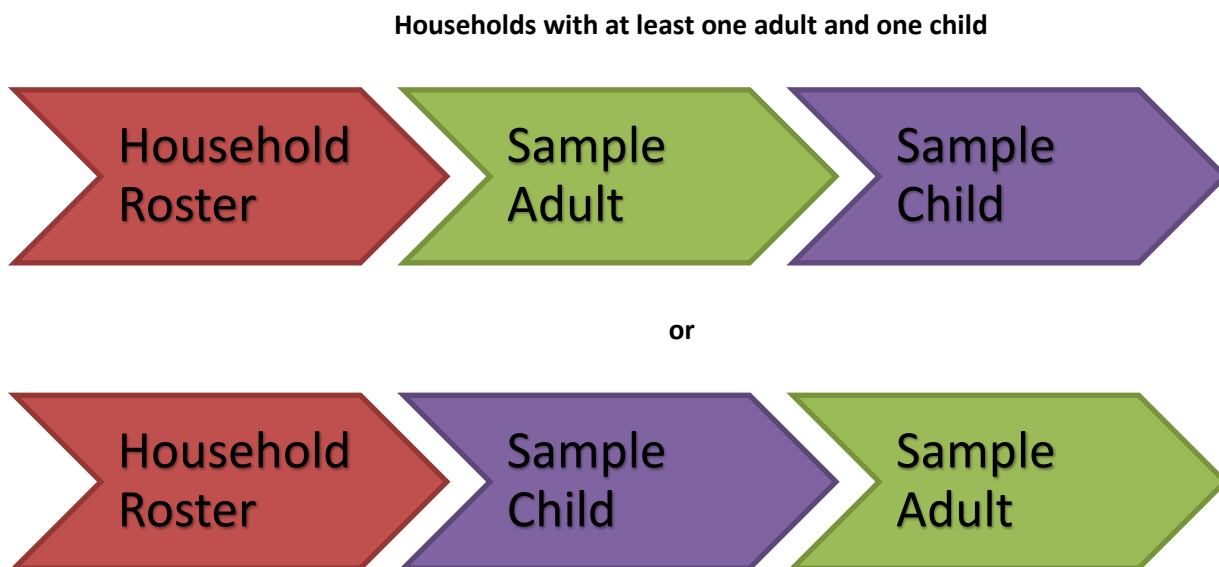
A portion of the redesigned questionnaire was fielded in June 2018 in select U.S. locations to pilot procedures and questions. In September 2018, a comprehensive two-day in-person training focusing on the 2019 redesign was offered to field interviewers and supervisors conducting NHIS interviews across the nation. During Quarter 4 (October–December) of 2018, the redesigned 2019 structure and content was fielded using a split-sample to provide practice to field staff and finalize operations and procedures including quality-control measures.

#### Structure of 2019 Redesigned NHIS

One “sample adult” aged 18 years or older and one “sample child” aged 17 years or younger (if any children live in the household) are randomly selected from each household following a brief household rostering interview that collects basic demographics of everyone who usually lives or stays in the household and identifies which members of the household are in the sample adult’s and sample child’s families. Information about the sample adult is collected from the sample adult him/herself unless s/he is physically or mentally unable to do so, in which case a knowledgeable proxy can answer for the sample adult. Information about the sample child is collected from a parent or adult who is knowledgeable and responsible for the health care of the sample child. This respondent may or may not also be the sample adult.

Figure 1 illustrates the new interviewing flow of the Household Roster, Sample Adult and Sample Child interviews, or modules, in the NHIS. The sample adult and sample child may be part of the same family or be part of different families in the household.

Figure 1. 2019 NHIS Redesigned Structure



Note. In households where there is an eligible sample adult and sample child, either the Sample Adult or Sample Child interview is administered first. Once both the Sample Adult and Sample Child interviews have been completed the interview is complete.

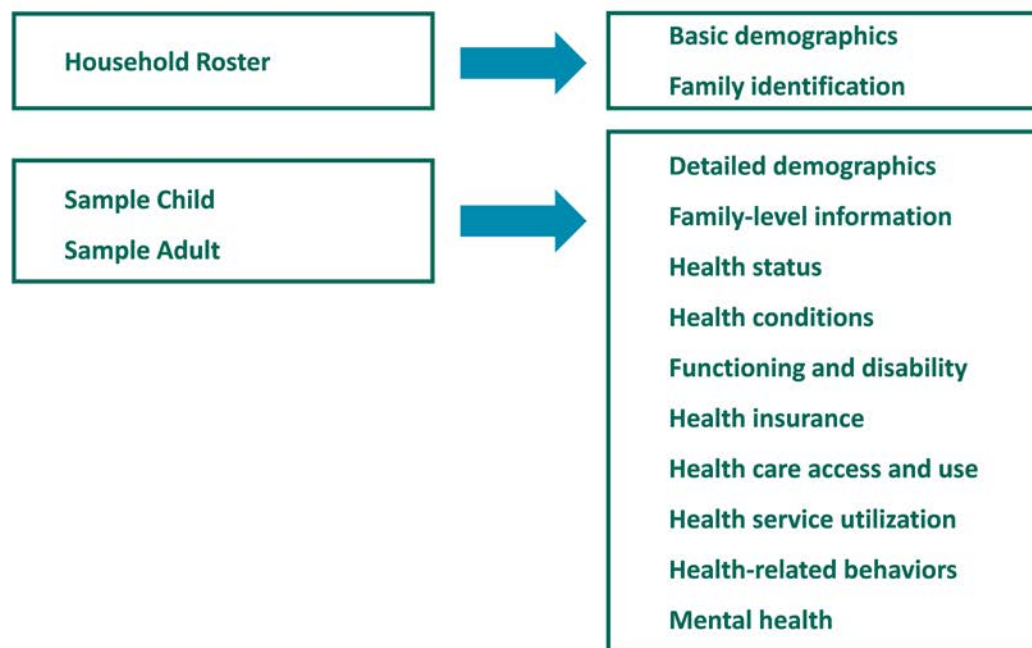


Note: The NHIS is a survey of the civilian noninstitutionalized population, and active duty military personnel are not included. In the rare case where a child lives in a household consisting of only active duty military personnel, the Household Roster would be completed followed by a Sample Child interview.

In the redesigned survey, family-level content is collected in the Sample Adult and/or Sample Child questionnaire module. Figure 2 illustrates how topics or content of the interview is organized by interviewing modules (i.e. Household Roster, Sample Adult and Sample Child). In contrast, the 1997–2018 NHIS administered questions about the family separately for each family in the household. The family interview asked about the family as a whole and about each member of the family. An adult family respondent provided information about him/herself and proxy information about the other family members. A sample adult and a sample child (if any children in the family) were randomly selected from each family to answer additional questions about their health. There is some loss of the content that has previously been collected in the family questionnaire, including: detailed relationships of all family members to household and family respondent, and information about each family member on various topics, including health insurance, access to health care and utilization of select health care services, country of birth (if not US), some employment and earnings information, active duty military time periods, disability-associated conditions, and receipt of medical advice by phone. For additional

information about the 1997–2018 content, refer to year-specific NHIS documentation:  
<https://www.cdc.gov/nchs/nhis/data-questionnaires-documentation.htm>

Figure 2. Topic organization in the Household Roster, Sample Adult and Sample Child modules: NHIS 2019.



### Content of the 2019 Redesigned Questionnaire

The redesigned NHIS questionnaire consists of three modules: (1) Household Roster; (2) Sample Adult Interview; and (3) Sample Child Interview.

#### *Household Roster*

In the Household Roster, an adult (aged 18 years and over) living in the household provides basic information about themselves and other people living in the household. The names, age, sex, race, and ethnicity of everyone in the household are collected. Additionally, the parents of all children are identified. The instrument then randomly selects one adult (sample adult) and one child (sample child), if any children live in the household, to be given follow-up questions. The sample adult is selected randomly among persons aged 18 years and over living in the household, and the sample child is selected randomly among those aged 17 years or younger. Questions are asked to determine who is in the family of the sample adult and sample child. The sample adult and sample child do not need to be in the same family. No health information is collected in this section. When the Household Roster is complete, the field interviewer can then proceed with the Sample Adult or Sample Child interview (if a child lives in the household).

---

## Sample Adult Interview

The sample adult is asked a series of health questions about themselves. Some questions vary by age or sex of the sample adult, but most are the same for all sample adults. Additional demographic information is also collected about the sample adult and his/her family.

## Sample Child Interview

An adult knowledgeable and responsible for the health of the child is asked a set of questions about the sample child. Some questions asked of the sample child vary by age, as younger and older children have different health needs. Additional demographic information is also collected about the child and his/her family.

## Types of Questions

The redesigned NHIS incorporates a long-term structure of fixed and periodic content. The long-term structure for the Sample Adult and Sample Child questionnaires organizes question topics by year and by type of content for the survey years 2019–2027. Additional information about periodicity of question topics for 2019–2027, see, [https://www.cdc.gov/nchs/nhis/2019\\_quest\\_redesign.htm](https://www.cdc.gov/nchs/nhis/2019_quest_redesign.htm). There are four types of content: (1) Annual core; (2) Rotating core; (3) Sponsored content; and (4) Emerging topics.

**Annual core** are consistent questions that are asked every year.

**Rotating core** are questions that are asked some, but not all years. These questions are scheduled to appear on a periodic basis of every other year, one out of every three years, or two out of every three years.

**Sponsored content** are questions funded by other federal agencies or other centers within CDC about topics of interest to the sponsor. Sustaining sponsors fund content every year, whereas other sponsors fund content periodically.

**Emerging topics** are questions about areas of interest to NCHS, CDC, or DHHS. These are newer subject areas that have generally not been researched in the general population.

## Questionnaire Sections

The NHIS is divided into many questionnaire sections within each module, each with a different focus. The sections may include any combination of annual core, rotating core, sponsored content, or emerging topics. When the same questions or same types of questions are asked in a Sample Adult and Sample Child interviews, the sections are given the same name for both interviews. The names of the questions asked of the sample adult or pertaining to the sample adult's family all end in “\_A” whereas those asked of the sample child or about the sample child's family end with “\_C.” Section names have a 3-letter abbreviation (e.g., INS for Health Insurance), and questions are grouped by module and section.

## Description of the 2019 Redesigned Questionnaire

A description of the 2019 topics and type of questions are described in this report under Sample Adult's Health, Sample Child's Health, and Health Insurance, while all demographic information has been portioned into four



sets of characteristics: 1) those about the sample adult and sample child; 2) those about the parents or guardian residing in the household with the sample child; 3) those about the spouse or partner residing in the household with the sample adult (if married or cohabiting); and 4) those about the family of the sample adult and sample child. In this document, multiple questionnaire sections may be described in each of the health topics included under Sample Adult's Health and Sample Child's Health.

Sample Adult health topics for 2019 are:

- I. Health Status and Conditions  
asthma, arthritis, cancer, cardiovascular conditions, diabetes, cholesterol, hypertension, and other chronic conditions, self-reported health status, pregnancy status and height and weight
- II. Functioning and Disability  
anxiety, communication, cognition, depression, hearing, mobility, self-care and upper body, social functioning (participation), and vision
- III. Pain and Pain Management  
chronic pain and opioid use
- IV. Health Care Access and Health Service Utilization  
difficulty paying for health care, utilization of services, immunizations, dental care, mental health care, physical and other therapeutic care, and prescription medication
- V. Health-Related Behaviors  
cigarettes, e-cigarettes, and other tobacco products
- VI. Mental Health  
generalized anxiety disorder (GAD-7) and depressive disorder (PHQ-8) scales
- VII. Preventive Care  
aspirin use and cancer screenings

Sample Child health topics for 2019 are:

- I. Health Status and Conditions  
health status, asthma, diabetes, developmental conditions, and learning disabilities
- II. Functioning and Disability  
anxiety, behavior, cognition, communication, depression, hearing, mobility, self-care and upper body, and vision
- III. Health Care Access and Health Service Utilization  
difficulty paying for health care, utilization of services, immunizations, dental care, mental health care, physical and other therapeutic care, and prescription medication
- IV. Behavioral and Mental health  
Baby Pediatric Symptom Checklist (BPSC) and Strengths and Difficulties Questionnaire (SDQ)
- V. Stressful Life Events



## IV. Sponsors

Some 2019 NHIS content is sponsored by other federal agencies or other centers within the Centers for Disease Control and Prevention. These sponsors purchase content that can be on the NHIS annually or periodically. Sponsored content may be used to collect data on new topics or to go into more depth about subjects already on the NHIS. Sustaining sponsors are agencies that sponsor content every year over multiple years.

In 2019, questions about cancer screening, immunization, exposure to health care settings, food security, Supplemental Nutritional Assistance Program (SNAP) assistance, non-cigarette tobacco use, arthritis, and insulin use were sponsored by the following federal agencies:

### Cancer Screening



The National Cancer Institute at the National Institutes for Health (NIH/NCI) and the National Center for Chronic Disease Prevention and Health Promotion at the Centers for Disease Control and Prevention (CDC/NCCDPHP) sponsored 25

questions asked of sample adults about screenings for breast, cervical, prostate, and colorectal cancers. Many of these questions provide more detailed information about topics already included in the 2019 NHIS rotating core.



### Immunization and exposure to health care settings



The National Center for Immunization and Respiratory Diseases at the Centers for Disease Control and Prevention (CDC/NCIRD) sponsored 17 sample adult questions about flu vaccination during pregnancy, pneumonia, tetanus, shingles, and HPV, and two sample adult questions about working or volunteering in the health care industry.

### Food Security and SNAP Assistance



The United States Department of Agriculture (USDA) sponsored 10 questions that can be used to determine food security or insecurity and degree of insecurity in the sample adult or sample child's family. Additionally, the USDA sponsors a question about use of SNAP in the past 30 days.

### Non-cigarette Tobacco Use



The Center for Tobacco Products at the Food and Drug Administration (FDA) sponsored seven sample adult questions about the use of cigars, pipes and smokeless tobacco.

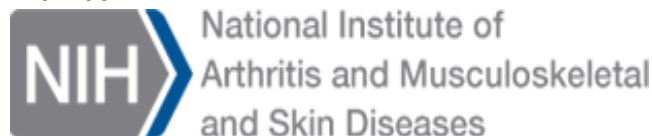
### Insulin Use



The National Institute of Diabetes and Digestive and Kidney Diseases at the National Institutes for Health (NIH/NIDDK) and the National Center for Chronic Disease Prevention and Health Promotion at the Centers for

Disease Control and Prevention (CDC/NCCDPHP) sponsored three sample adult questions about insulin initiation among adults with diabetes who take insulin.



**Arthritis**

The National Institute of Arthritis and Musculoskeletal and Skin Diseases at the National Institute of Health (NIH/NIAMSD) sponsored six sample adult questions about joint pain and arthritis, how it impacts the work of adults with arthritis, and if they have been told to lose weight or exercise to improve their pain.

## V. Sample Sizes and Response Rates

The NHIS sample size can vary from year to year. The normal annual sample size (i.e., the number of households for whom data are collected and publicly released) for the sample design starting in 2016 (and similar to the sample design covering 2006-2015) is about 35,000 households containing about 27,000 sample adults and 9,000 sample children. The sample size can be reduced for budgetary reasons or increased because supplementary funding is available.

A primary goal for 2019 was 30,000 completed Sample Adult interviews. The sample for 2019 was expanded by about 13 percent from 2018 to meet this goal, but the response rate in the first half of the year was better than expected and would have yielded considerably more than 30,000 Sample Adult interviews. As a result, a 10 percent reduction in the latter half of 2019 was implemented.

The publicly released data files (also called “public use data files”) for the 2019 NHIS contain data for 31,997 sample adults and 9,193 sample children from 33,138 households. In 568 cases, a knowledgeable proxy answered for the sample adult because he/she was mentally or physically incapable of answering for himself/herself.

### Response Rate Method

Response rates presented here conform to the American Association of Public Opinion Research (AAPOR) Response Rate Definition # 2, or AAPOR RR2 (AAPOR, 2016). “Interviewed households,” “interviewed sample adults,” and “interviewed sample children” include those with completed interviews or acceptable “sufficient partial” interviews.

In the NHIS, a sample adult or sample child interview is considered fully complete when respondents complete all sections, and therefore questions, for which they are eligible. Conversely, an interview is considered a “partial” when all sections are not completed. The most common reason for a partial is a “break-off,” which occurs when a respondent stops the interview in-progress before completion and the interviewer fails to complete the interview during the allotted assignment period. The partial interview rate is the percent of all sufficiently complete interviews that are not entirely complete. These “sufficient partials” are counted as interviews in the computation of response rates. Partial interviews that are not far enough along in the interview, known as “insufficient partials,” are considered refusals (Stussman et al., 2003) and therefore included as eligible, non-respondents in the computation of response rates.

### Household Response Rate

The household response rate is calculated by dividing the number of interviewed households (33,138) by the sum of the number of interviewed households (33,138) and the number of nonresponding households (21,093). Nonresponding households are eligible households that were not interviewed for a variety of reasons, including

language barriers, no one home after repeated contact attempts, refusal, household records rejected for insufficient data, or other reasons for no interview.

The total Household Response Rate for 2019 was 61.1%.

It is important to note that the definition of an interviewed household differs from the past design (1997-2018). Previously, an interviewed household was defined as one where at least one family in the household completed a substantial portion of the family interview. With the family interview removed from the redesigned NHIS, an interviewed household is now defined as one where the household roster and a substantial portion of either the Sample Adult interview or the Sample Child interview (if one or more children reside in the household) is completed. A household response rate obtained during the 1997-2018 NHIS and 2019 NHIS should be presented separately.

### ***Household Roster Completion***

The completion of the Household Roster is defined as the enumeration of all the persons in an eligible household, with basic demographic information collected about each household member. The Household Roster Completion Rate is calculated by dividing the number of eligible households with a completed household roster (35,404) by the number of eligible households (54,231). For 2019, the Household Roster Completion Rate was 65.3%. Based on demographic information obtained from completed household rosters, there were 10,155 eligible sample children and 35,365 eligible sample adults.

### ***Sample Child Response Rates***

Sample child response rates can be computed two ways, resulting in either a conditional or final response rate. The Conditional Sample Child Response Rate is calculated by dividing the number of interviewed sample children (9,193) by the number of eligible sample children from households with completed rosters (10,155).

For 2019, the Conditional Sample Child Response Rate was 90.5%.

The Final Sample Child Response Rate accounts for the Household Roster Completion Rate and is calculated by dividing the number of interviewed sample children (9,193) by the number of eligible sample children (10,155) from households with completed rosters, and then multiplying this quotient by the Household Roster Completion Rate (65.3%). In 2019, 0.9% of Sample Child interviews were sufficient partials.

For 2019, the Final Sample Child Response Rate was 59.1%.

### ***Sample Adult Response Rates***

As with sample children, both a conditional and final response rate can be computed for sample adults. The Conditional Sample Adult Response Rate is calculated by dividing the number of interviewed sample adults (31,997) by the number of eligible sample adults from households with completed rosters (35,365).

For 2019, the Conditional Sample Adult Response Rate was 90.5%.

The Final Sample Adult Response Rate is calculated by dividing the number of interviewed sample adults (31,997) by the number of eligible sample adults from households with completed rosters (35,365), and then

multiplying this quotient by the Household Roster Completion Rate (65.3%). In 2019, 3.1% of Sample Adult interviews were sufficient partials.

For 2019, the Final Sample Adult Response Rate was 59.1%.

Note that numbers of households, sample children, and sample adults eligible and interviewed were used for the calculations of response rates shown and rounding discrepancies may occur when using the percentages.

### Reporting Household, Sample Adult, and Sample Child Response Rates

Which response rate to report depends on the focus of one's analysis. When reporting on analyses performed with the Sample Adult data file, the data user should report the Final Sample Adult Response Rate. In addition, it is good practice to also report the Household and Conditional Sample Adult Response Rates. Similarly, if the focus of one's analysis is the sample child, the Final Sample Child Response Rate should be reported. Again, the Household and Conditional Sample Child Response Rates could also be reported.

The Household Response Rate would also be reported if one is performing a household-level analysis using the public-use paradata file.

## VI. Weighting

NHIS is a sample survey. That is, only a sample (subset) of the civilian noninstitutionalized population is selected to participate in the survey. Additionally, not everyone selected to participate agrees to participate, which can affect the representativeness of the sample. In order to account for these two factors, sampling weights are created. These sampling weights are used to produce representative national estimates. The data must be weighted to obtain population estimates for survey outcomes in the population represented by the NHIS. The value of the weight for a given respondent can be interpreted as the number of persons in the NHIS target population represented by that respondent. The sum of the weights over all respondents is used to estimate the size of the total target population. The weights reflect several steps of adjustments starting with a base weight, which is inverse to the probability of selection. Households and persons that are more likely to be selected are given lower weights so that the final estimates are not biased by their increased likelihood of being selected. For example, in a household of two eligible adults, the sample adult has a selection probability of one-half, and therefore their base weight will be increased by two. However, in a household of four eligible adults, the sample adult has a selection probability of one-fourth, and therefore their base weight will be increased by four, since roughly speaking they represent more people from the household. The base weights are then adjusted for nonresponse patterns, that is, the different response rates among different household and person-level subgroups.

The 2019 questionnaire redesign provided an opportunity to evaluate the adjustment approach that had been in place since 1997. For 1997-2018, the adjustment approach was based on geography; the weights for households and persons in geographic areas with lower response rates were increased more than for those in areas with higher response rates. That way, final estimates were not biased by the latter group's increased likelihood of participation. More sophisticated methods to decrease potential nonresponse bias are now available (Olson, 2013; Valiant et al., 2018), and based on the evaluation, the weighting process for 2019 was updated. The updated approach for nonresponse adjustment uses multilevel regression models that include paradata variables that are predictive of both survey response and selected key health outcomes, the key criteria for effective bias reduction.

Finally, the nonresponse adjusted weights are calibrated to U.S. Census Bureau population projections and American Community Survey (ACS) one-year estimates for age, sex, race and ethnicity, educational attainment, Census division, and Metropolitan Statistical Area (MSA) status. Prior to 2019, calibration was only to age, sex, and race and ethnicity population projections. These changes to the nonresponse adjustment approach and the calibration methods have the potential to impact comparisons of the weighted survey estimates over time.

A report with further information about NHIS sampling weights is available on the 2019 data release page at <https://www.cdc.gov/nchs/nhis/2019nhis.htm>.

For 140 households in 2019, there were inconsistencies in the data that suggested the household roster was incomplete. The number of adults and number of children in these households is unknown, and the probabilities of selection of the sample adult and sample child in these households were based on the incomplete roster. This had the potential to introduce bias, and the potential level of bias will be evaluated for key estimates, but it is expected to be negligible.

## VII. Variance Estimation

In a data collection, estimates based on different samples will vary and can differ from the true population values. The estimated difference between the true target population value and the estimate from a random sample is the sampling error. Sampling error cannot be directly calculated because the true target population value is unknown. Rather, sampling error is estimated and expressed as a standard error (SE), the average degree to which estimates based on random samples differ from each other and the true target population value due to sampling. This measure is incorporated in common statistical methods such as significance testing and estimating confidence intervals.

Because of the complex nature of the NHIS sampling design (specifically, the use of stratified cluster sampling), key nesting variables were created to capture explicit stratification and to identify clustering for a more accurate estimation of the sampling error. For additional information about using weights and variance variables in analysis, including examples in selected statistical software packages, see “Using weights and variance estimation variables, and examples.”

Analysts should be aware that the use of standard statistical procedures that are based on the assumption that data are generated via simple random sampling (SRS), instead of a complex sample design, generally will produce incorrect estimates of variances and standard errors when used to analyze data from the NHIS. Analysts who apply SRS techniques to NHIS data generally will produce standard error estimates that are, on average, too small, and are likely to produce results that are subject to excessive Type I error.

## VIII. Editing the Data During and After the Interview

### Edits to Protect Confidentiality

NCHS (including its contractors and agents) collects personally identifiable NHIS and other survey data under a pledge of confidentiality and a promise that the data will be used only for statistical purposes. Section 308d of the Public Health Service Act and Section 512b of the Confidential Information Protection and Statistical Efficiency Act (CIPSEA) require that confidentiality be maintained without exception. Violations of CIPSEA are a class E felony, punishable by imprisonment for not more than 5 years, a fine of not more than \$250,000, or both.

Strict procedures in survey operations and data dissemination are used by NCHS, its data collection contractors, and other agents to prevent disclosure of survey subjects' identities.

The risk of inadvertent disclosure of confidential information regarding individual respondents is higher when there exists a publicly released data set having detailed geography variables and a detailed and extensive set of survey observations. For this reason, the NHIS does not publicly release state identifiers and some other geographic variables, and the original design strata and primary sampling units (PSUs) are masked when the data are publicly released. NHIS data sets may also be coarsened by suppressing survey variables, collapsing multiple variables into one, and collapsing response categories. In addition, statistical noise at both the variable level and record level may occasionally be added to protect confidentiality.

Notes fields in the Codebook report may include information about edits and data suppression that were done to protect the confidentiality of NHIS participants. However, one important edit is worth noting here because it applies to multiple variables across the survey. To protect confidentiality among the oldest adults, all age variables were top-coded to "85 years and older" (85+). For example, survey questions related to age at diagnosis for cancer (i.e. LUNGAGE\_A) and diabetes (DIBAGE\_A) ("How old were you when you were diagnosed [with this condition]?") are top-coded to 85+ years. The recode DIFIRSTC\_A ("Years since first diagnosed with diabetes") is also top-coded to 85+ years to ensure confidentiality.

To further protect confidentiality, detailed information for some variables are not available on the NHIS public use data files. For a list of questions not available on the public use file, see the restricted-use codebook available on the 2019 data release webpage, and the Appendix in this document.

Analysts interested in working with data that were suppressed or edited to protect confidentiality may apply to access selected unmodified data files through the NCHS Research Data Center (RDC). The RDC is a data enclave established to provide a mechanism whereby researchers can access detailed data files in a secure environment without jeopardizing the confidentiality of survey participants. Information about RDC access options and application procedures is available at: <https://www.cdc.gov/rdc/>.

## Family-Level Replicate

In the field, the interviewer can conduct either the Sample Adult or Sample Child interview first in households where both eligible adults and children reside. In instances where the sample adult and the sample child belong to the same family, the instrument is optimized to only ask family level questions in the first interview. This helps to minimize respondent burden by eliminating repetition for family level questions, such as family income and food security. If, however, the respondent of the first interview refuses or doesn't know the answer to a significant number of questions within a family-level section, that section is repeated in the second interview when the respondent of the second interview isn't the same individual.

The family level data collected are then replicated (i.e., copied) to the other interview to a replicate variable of the same name (but a different suffix) in a post-processing step. For example, if the adult interview preceded the child interview and they are in the same family, the question about whether anyone in the family had problems paying medical bills is collected in the adult variable PAYBLL12M\_A and replicated to the child variable PAYBLL12M\_C. The Questionnaire report identifies a variable as being replicated in the "Replicate to:" field. Searching the document for the string "Replicate" will identify the variables that underwent replication.

When the sample adult and the sample child are in different families within the household, both the sample adult and the sample child respondent will be asked family level questions about their respective families. In



households where there are no children or there are no eligible adults (e.g. all active Armed Forces), there is no replication involved.

Annual core sections of the questionnaire with instrument optimizations and replicated variables include Family Income (INC), Family Employment (FEM), Difficulty Paying for Health Care (PAY), Food-Related Programs (FOO), Housing (HOU), and Telephone Use (TEL). Depending on the rotation structure for the year in question, they can also be found in some sponsored sections of the questionnaire such as Food Security (FDS) and Food-Related Programs (FOO).

## Hard and Soft Edits

To help prevent both interviewer data entry error and respondent error, range values and consistency checks may be programmed into the CAPI system. During the interview, if an interviewer enters an out-of-range value (such as 180 years instead of 18 for age), an error message instructs the interviewer to enter a new value. Such an interruption of the interview is called a “hard edit” if the interview cannot continue without an acceptable response being entered, and a “soft edit” if the interview may continue with or without a new response being entered. Soft edits may apply to questions for which the response entered is plausible (such as an extreme height value).

Even with such checks built into the CAPI system, data cleaning (data “editing”) is still necessary. The first step in the data cleaning process is verification of the valid number of cases in the data file and the review of frequencies for reasonableness. Each variable is examined to determine if its values are within its range of permissible values. Values not in that range are verified as missing if they are not in the universe due to legitimate skip patterns in the questionnaire or set to the special value of “not ascertained” if there was a break-off in the interview.

## Question-Specific Replication

An optimization edit is an edit that fills-in values for variables that were skipped in the instrument because the information could be inferred from the Sample Adult or Sample Child interview, whichever went first. For example, the marital status of the sample adult would be known if the Sample Child interview preceded the Sample Adult interview, if the sample adult and sample child were in the same family, and if the sample adult was also a parent of the sample child. Select questions in health insurance relating to detailed characteristics of shared private plans between the sample child and sample adult in the same family were also filled-in from responses of the interview that came first.

## Recode into Different Variable

Recodes have been created for select questions to make the data more analytically useful. One example of this is a recode that converts a single variable allowing selection of as many answers as are applicable into a series of variables (one for each possible response) with yes/no or mentioned/not mentioned responses. Other recodes have been created to summarize information obtained from multiple questions available to the public (e.g. summary scores of validated scales), or to combine information from which some information may not be publicly available (e.g. multiple race categories).

Data users are recommended to review the description of ‘Major Recodes’ for the topic of interest in this document, and the codebook documentation for additional recode information.

---

## Orientation to How to Use NHIS data

### I. Survey Data Files and Documentation

All datasets and associated documentation for 2019 are available on the NHIS website,

<https://www.cdc.gov/nchs/nhis/2019nhis.htm>

Documents in Portable Data Format (PDF) describing NHIS methods, survey implementation materials and other background information are included under the tabs “Using the NHIS” and “Survey Implementation Materials.”

The following documents are included in the tab “**Using the NHIS:**”

**Survey Description document (PDF):** A description of NHIS methods, year-specific response rates and content, and other useful resources for NHIS data users.

**Weighting guidance (PDF):** A description of the procedures for adjusting NHIS sampling weights to account for nonresponse to the survey in the 2019 redesigned questionnaire.

**Imputed income technical document (PDF):** A description of the methodology for creating the 2019 NHIS imputed income variables.

**File record length and size summary (PDF):** A year-specific list summarizing the number of records, file size, and record length for each of the ASCII data files released.

**Paradata Survey Description document (PDF):** A year-specific description of the interview process information collected from sampled households.

**Notice for data users (PDF):** A year-specific log documenting data releases and other file updates.

**Checksum (PDF):** A list of year-specific reference values for each ASCII and CSV data file released to allow data users verify the integrity of downloaded files.

The following documents are included in the tab “**Survey Implementation Materials:**”

**Survey Questionnaire - English (PDF):** Year-specific NHIS questions fielded. A new format was introduced in 2019 and described below.

**Survey Questionnaire - Spanish (PDF):** Spanish version of the year-specific NHIS questions fielded.

**Field Representative Manual (PDF):** The manual on Computer-Assisted Personal Interview (CAPI) for NHIS interviewers.

**NHIS Instrument Flowchart (PDF):** A graphical view of the questionnaire content.



**NHIS Sponsored Content (PDF):** A year-specific sponsoring agencies and associated sponsored questions in NHIS. A new format was introduced in 2019, described below, that replaces the NHIS Sponsors and Supplements Report available during 1997-2018.

**Survey Brochure (PDF):** A tri-fold sheet that describes the NHIS to survey participants.

**Advance Letter (PDF):** A letter that explains the NHIS mailed to sampled households prior to interviewer contact and provided to survey respondent at the time of the interviewers' visit.

## Survey Questionnaire

The 2019 survey questionnaire (PDF) lists the questions in the survey and descriptive information about them. The information in the document is organized in two panels: a hierarchical bookmarks panel on the left for navigation, and a main panel on the right for displaying detailed content. The bookmarks themselves are organized as:

- a) Link to a contents page that explains the document's overall structure
- b) Hierarchical section index that lists the sections, their descriptions, the content type (Annual Core, Rotating Core, Sponsored Content or Emerging Content) and the page range in the PDF for each section for ease of printing
- c) Hierarchical questionnaire organized by module, section and variable and appearing in the order that the questions are asked.

Selecting a bookmark for a module or a section navigates to the first variable in the module or section, respectively. When selecting a variable of interest, detailed information for that variable is displayed in the main panel. For each variable, the main panel heading has the year and title of the survey along with the section abbreviation and description. The body of the main pane starts with a header (in blue background) with the Question ID (used for ordering questions in the questionnaire), the variable name, the interview module and the content type. This is followed by the question text as it appears in the instrument. The question text may contain one or more context-sensitive fills, indicated by text with a leading caret (^) symbol. A fill is text that is conditionally generated to modify the question text to make it more suitable to the context of the interview. For example, the fill whose name is ^heshe\_C will expand to "he" if the sample child is male, "she" if female or "they" if sex is not known. All fills in the question text appear in the fills table in order along with their description and rule-based instructions on how the fill text is generated in the instrument. For some variables, as part of the question text, there may be interviewer instructions in bolded blue text with any optional text appearing in italics and gray font. Below the fills table, there is another table with valid response categories and their descriptions, followed by the universe description, and if present, any skip instructions, hard or soft edits.

Both English and Spanish versions of the questionnaire are available on the NHIS website. The Spanish version of the questionnaire has Spanish translations for the question text, the fills and the response choices.

## NHIS Sponsored Content

The NHIS Sponsored Content (PDF) lists the sponsoring organizations and the questions that they have sponsored in the NHIS for the current year. The information in this document is organized in two panels: a bookmarks panel on the left listing the sponsoring organization names alphabetically (short form), and a main

panel on the right listing the variables and some descriptive information about them such as question ID, question text and the universe description. When content is sponsored by multiple organizations, the names of all the sponsors are shown on the bookmarks panel. Selecting an organization's bookmark takes you to the content sponsored by that organization. The document's page header lists the sponsoring organizations' full name(s).

## Data Files

The 2019 NHIS data release includes files for the Sample Adult, Sample Child, Imputed Income for the sample adult and sample child, and Paradata. The file names for the 2019 data release are listed in Table 1. Files corresponding to the 2019 NHIS have a two-digit suffix at the end that represents the survey year, e.g., 19 for 2019, or adult19. In prior years, separate files were available for household, family, and person level information due to the different survey design. Imputed income files for sample adult and sample child can be merged with their respective Sample Adult and Sample Child files to create a single dataset. Sample Adult and Sample Child files can also be merged with Paradata.

### *Sample Adult and Sample Child files*

The 2019 Sample Adult and Sample Child files include all publicly available questionnaire variables and associated recodes, and household and family-level variables. The Codebook and Summary reports for each file describe their contents in detail.

### *Imputed Income Files*

The 2019 Imputed Income files for sample adult and sample child contain 10 imputations of family income and poverty ratio as both continuous and categorical top-coded variables. An example with sample code that demonstrates using the imputed income data file in an analysis is described in this report in the section "Merging Files" under the heading "Using Imputed Income Data Files."

Variables from the first imputation were also added to the Sample Adult and Sample Child files for convenience of users who do not use multiple imputed data in their analyses. While each of the 10 imputations has been drawn from a valid distribution based on a regression model, the first imputation included in the Sample Adult and Sample Child files may be slightly different to the other set of imputations. Single imputation analyses result in estimated standard errors that are too small because the imputed values are treated as if they were observed. This ignores the inherent uncertainty resulting from lack of knowledge about the true (unobserved) value, but it is superior to analyses that use only cases with observed values.

Information about income measurements in NHIS and income recodes are described in this document in the sections "Family Income" and "Recodes of Family Income and Imputed Family Income." Methodology for creating the imputed income variables are described in the "Imputed Income Technical Document" available with the 2019 file releases on the NHIS website, under "Using the NHIS."

### *Paradata File*

The NHIS Paradata File contains information about the interview process. The data from the Paradata File are collected as part of the NHIS interview, using computer-assisted personal interviewing (CAPI). The NHIS paradata come from a number of sources:

- The Contact History Instrument (CHI), a supplemental piece to the NHIS that collects data from the interviewer about each contact attempt. Data include strategies used for gaining participation and reasons for respondent reluctance.
- The Back section of the NHIS, where a series of questions are asked of the interviewer, including mode of interview (in-person visit vs. phone interview), and reasons for partial interviews/breakoffs.
- Date and time variables from each module of the instrument (Household, Sample Child, Sample Adult). The date and time information are collected each time a module is started and completed.

The Paradata File is on a case (household) level, where one record represents one case. Unlike the NHIS public use Sample Adult and Sample Child data release, which contains information on fully complete and sufficiently complete interviewed cases only, the Paradata File also contains data on other types of cases, including cases that were ultimately refusals, insufficient partials, and other types of nonresponse. These are referred to as *Type A* cases. In addition, data are provided on cases that were deemed out of scope, such as families with Armed Forces-only adults. The out-of-scope cases are referred to as *Type B* outcomes in this document.

The Paradata File is intended as both a stand-alone data file and one whose fully complete and sufficiently complete cases can be linked with the Sample Adult and/or Sample Child data files. For more information about the paradata file, including linking paradata files with other data files, see the Paradata Survey description Document available with the 2019 file releases on the NHIS website, under “Using the NHIS.”

### ***File Names***

The Sample Adult, Sample Child, Imputed Income and Paradata files are released as both a column-delimited text (ASCII) file and a comma-separated values (CSV) file. Programs that contain input statements in SAS, Stata and SPSS environments are provided to help load the ASCII files into datasets with proper data formats and labels in the respective computing environments. Table 1 list the names of data files, programs and documentation in the 2019 NHIS data release.

### **Data Documentation**

Summary and Codebook PDFs provide accompanying documentation for the 2019 Sample Adult, Sample Child, Imputed Income and Paradata files.

### ***Summary (PDF)***

The summary report is a PDF document with a bookmarks panel on the left organized by module and section in questionnaire order, and a main panel that displays the variable list by section. When the section is selected in the bookmarks panel, the following information for all variables in that section is displayed in the main panel in tabular form. The lead-in header has the module name, the 3-letter section abbreviation and the section description. The table has rows with the following information:

**Question ID:** if the variable is in the questionnaire, the unique ID for that variable is displayed. For recodes, the word “Recode” is displayed, and for any that are neither (e.g., identifiers), this column is blank. Question ID may change by survey year and should not be used for data management purposes across years.

**Variable name:** the name of the variable in the data. Generally, the variable name in the questionnaire has the same variable name in the dataset.

**Source variables:** for recodes, this column lists the names of variables used to create the recode

**Description:** the variable label

**Type:** the data type for this variable, i.e., character or numeric

**Location:** the column range in the ASCII file (column numbers) where this variable is stored

**Length:** the length of the variable as a character data type

Table 1. Data release files names for the Sample Adult, Sample Child, Imputed Income and Paradata files: NHIS 2019.

Type of file	Sample Adult files	Sample Child files	Imputed income files	Paradata files
Data in column-delimited ASCII format	adult19.dat	child19.dat	Adultinc19.dat Childinc19.dat	paradata19.dat
Data in comma separated Values (CSV)	adult19.csv	child19.csv	Adultinc19.csv Childinc19.csv	paradata19.csv
SAS program with input statements	adult.sas	child.sas	adultinc.sas childinc.sas	paradata.sas
STATA program with input statements	adult.do	child.do	adultinc.do childinc.do	paradata.do
SPSS program with input statements	adult.sps	child.sps	adultinc.sps childinc.sps	paradata.sps
Summary	Adult_summary.pdf	Child_summary.pdf	Adultinc_summary Childinc_summary	Paradata_summary.pdf
Codebook	Adult_codebook.pdf	Child_codebook.pdf	Variables included in the Adult and Child Codebooks	Paradata_codebook.pdf

### Codebook (PDF)

The Codebook report combines all the detailed information for a variable with the unweighted frequencies (counts and percentages) found in the data. The Codebook report is a combination of the former variable layout and variable frequency reports from prior years. The report has a navigational bookmarks panel on the left with expandable module and section bookmarks in questionnaire order. The main panel on the right contains the

variable detail. When a variable bookmark is selected, the detailed display includes its module, section, file, data type, question text (if present), question fill information, universe and universe description, the variable description or label, question ID, keywords, and notes.

This is followed by a table that provides the unweighted frequencies and percentages for the variable. All response categories are shown in the table, including those with a zero count in the data files. For continuous variables, a range of values is provided. This allows users to see a complete list of response categories with frequencies for each variable without referring to additional documentation. In addition, the “frequency missing” label will be shown if a variable has cases that are not in the universe.

In the NHIS, the same codes are used across all files to designate “refused” (RF) and “don’t know” (DK) responses: refusals are coded as 7 (with leading 9’s added to the length of the field, as in 7, 97, 997, etc.), while “don’t know” responses are coded as 9 (with leading 9’s added to the length of the field, as in 9, 99, 999, etc.). For partially completed interviews (e.g., Sample Adult interviews where the respondent discontinued the interview before reaching the question), the responses will appear as 8’s for “not ascertained”, again with leading 9’s added to the length of the field, as in 8, 98, 998, etc., for the remaining variables in the file. A code of 8 is also used to indicate “not ascertained” responses when the field was blank or contained an impossible code. Lastly, in some limited situations (primarily recodes), the “Refused,” “Don’t know,” and “Not ascertained” categories may be collapsed into a single category called “Unknown,” which is typically designated with a 9 (with leading 9’s to fill out the field, if necessary). Data users are advised to read the notes in the data release documentation for further information about the variables of interest.

### *Codebook for restricted-use variables (PDF)*

This document lists the restricted-use (or inhouse) variables that are available to analysts in the Research Data Center (RDC). It does not include any variables that are in the public data files. The format is similar to the codebook, except that no frequencies are shown.

## Variable Conventions

Variable labels are restricted to 80 characters due to limits in some programming languages. All variables have a length of 12 characters or less. Variables names in the Sample Adult file have the suffix `_A`, e.g., `DIBEV_A`, to indicate that they refer to the sample adult or were asked of the sample adult’s family. Variables associated with the sample child will analogously have the suffix `_C`, e.g., `DIBEV_C`. Variables that do not have these suffixes are household or family level variables or identifiers, e.g., `HHX`.

## Definitions

The following defines some terms used in the different reports:

**Fills:** Text that modifies the question, based on previously collected information and using conditional logic. Fills are indicated by a caret (^) symbol followed by the name of the fill, e.g., `^SCNAME`.

**Keywords:** Descriptive words or phrases relevant to the topic of the variable; these can be used for word searches.

**Notes:** Additional information that analysts need to know about a variable, such as assumptions, limitations, caveats, and differences between instrument versions. Analysts are encouraged to read the notes pertaining to variables of interest. Notes may contain cross-references to other pertinent variables.

**Recode:** A variable derived from the reordering, collapsing, or verbatim coding of another variable. Alternatively, a recode may be constructed from two or more variables. All variables used to construct a recode are listed as a cross reference in Sources. Users will note that several standardized variables appear in the NHIS dataset. A standardized variable is a type of recode based on time unit information obtained during the interview. When respondents are asked questions pertaining to time—for example, how long the respondent has worked at his/her job—the answer is typically obtained in two parts: the number of time units and the type of time unit. During data editing, this information is standardized into a single appropriate time unit. Some of the standardized time unit recodes may also be top-coded (i.e., the maximum reported value may be capped) for confidentiality reasons.

**Sources:** If the variable is a recode, then all variables that were used to make that recode are listed as sources.

**Universe:** The group of adults or children to whom a specific question applies. For example, the universes for most sample adult variables are adults who were age 18 or over. This universe is specified on the Codebook report as HHSTAT\_A=1. Sample adults who are not eligible to answer a given question are considered to be not-in-universe. For example, sample adults who reported that they never had high cholesterol, e.g., CHLEV\_A having a value of 2, or RF or DK the response would not be eligible for a follow-up question CHL12M\_A about whether they had high cholesterol in the past 12 months. Universes for many questions are often age specific. In the redesigned NHIS, missingness in the sample adult or sample child's age is possible, and in those few cases the individuals would be ineligible for the question. Note that during rostering, when a person's age is not known, there are age-related follow-up questions to get at whether they are adults or children so the sample adult or sample child selections can be made. If the age is still not known, the interview terminates. Similarly, the sex variable (SEX\_A or SEX\_C) also allows for missing values, but the interview can proceed. Sex-specific questions for the sample adult and the sample child are not in universe when sex is unknown.

## II. Using Weights and Variance Estimation Variables

**To appropriately analyze NHIS data, it is necessary to utilize weights and variance estimation variables.** This is because the NHIS uses a complex sample design involving stratification and clustering designed to represent the civilian noninstitutionalized population of the United States and not all sampled respondents respond. If data are not weighted, severely biased estimates may result, such as producing estimates that are not representative of the NHIS target population. If the correct variance estimation variables are not used, then estimates of precision, such as standard errors, will likely be smaller than they should be. This will make the data appear to be more precise and will result in more statistically significant differences between estimates and in other analyses that are subject to excessive Type I error (rejection of a true null hypothesis).

### Weight Variables

The Sample Adult and the Sample Child file each have a unique and separate set of weights and variance estimation variables. Prior to the 2019 redesign, the NHIS included person- and household-level weights that could be used to make estimates based on data from all family members or for the household itself. But given the questionnaire redesign these weights are no longer applicable to the NHIS.

The two sets of weights included in each file are the Final Annual Weight and Interim Annual Weight.

The Final Annual Weight can be used to generate national estimates. This weight includes the design, ratio, nonresponse and calibration adjustments. The Paradata file does not include a final annual weight.

**The final annual weight in the Sample Adult file is WTFA\_A**

**The final annual weight in the Sample Child file is WTFA\_C**

The Interim Annual Weight is required by some software packages for certain methods of variance estimation. These weights do not include the calibration adjustment (age-sex-race/ethnicity-education-MSA status-Census division raking to population control totals).

**The Interim Annual Weight in the Sample Adult file is WTIA\_A**

**The Interim Annual Weight in the Sample Child file is WTIA\_C**

**The Interim Annual Weight in the Paradata file is WTIA\_PD.**

For analyses of quarterly data, the Sample Adult and Sample Child weights should be multiplied by 4 so that the sum of weights represents the associated U.S. population totals for the year.

## Variance Estimation Variables

Weights and the complex survey design must be taken into account for appropriate analysis and variance estimation. As described earlier in the Sample Design section, the data collected in the NHIS are obtained through a complex sample design that involves stratification and clustering. The use of standard statistical procedures that are based on the assumption that data are generated via simple random sampling (SRS) will produce incorrect estimates of variances and standard errors that are, on average, too small, and are likely to produce results that are subject to excessive Type I error when used to analyze data from the NHIS.

The Sample Child and Sample Adult public use files contain pseudo-design variables necessary for variance estimation using Taylor series linearization methods.

**For both the Sample Adult and the Sample Child files, the stratum and primary sampling unit (PSU) variable names are PSTRAT and PPSU.**

PSTRAT and PPSU are simplified versions of the true NHIS sample design variables created for the public-use files in order to protect the identity of survey respondents. The strata identifier is not directly related to state or density strata. When using the publicly available data files for estimation purposes, strata and PSU identifiers provided by NCHS are required to properly estimate variances. The use of these publicly available variance estimation variables may provide slightly different standard errors than the use of the confidential variance estimation variables used by analysts at NCHS. Data users who want access to the confidential variance estimation variables used by analysts at NCHS may apply to the NCHS RDC: <https://www.cdc.gov/rdc/>.

## Example: Applying Sample Weight and Variance Estimation Variables in Analysis

Several software packages are available for analyzing complex samples. Below are examples of computer code for specifying sample weight and variance estimation variables for standard error calculation code of means, percentages and totals with the NHIS data using SUDAAN, Stata, SPSS, SAS, and R software packages for illustrative purposes.



The limited public release design information requires a mathematical simplification that the PSUs be treated as if they were sampled with replacement (WR). The simplified design structure can be specified for the Sample Adult file with the following statements in selected software packages.

### *Example using SUDAAN*

```
PROC <DESCRIPT, CROSSTAB, ...> ... DESIGN = WR;
NEST PSTRAT PPSU;
WEIGHT WTFA_A;
```

Note that SUDAAN requires that the input file be sorted by the variables listed on the NEST statement (i.e., PSTRAT and PPSU). Design statements for other data files should use the appropriate weight variables found on these files.

### *Example using STATA*

```
Stata svy
svyset [pweight=wtfa_a], strata(pstrat) psu(ppsu)
svy: mean <name of variable to be analyzed for average>
or
svy: proportion <name of variable to be analyzed for percentage/proportion>
```

### *Example using SPSS*

SPSS csdescriptives (for averages) or cstabulate (for percentages/proportions):  
One needs first to define a “plan file” with information about the weight and variance estimation, e.g.:

```
CSPLAN ANALYSIS
/PLAN FILE="< file name >"
/PLANVARS ANALYSISWEIGHT=WTFA_A
/DESIGN STRATA=PSTRAT CLUSTER=PPSU
/ESTIMATOR TYPE=WR.
```

and then refer to the plan file when using csdescriptives or cstabulate, e.g.:

```
CSDESCRIPTIVES
/PLAN FILE="< file name >"
/SUMMARY VARIABLES =<name of variable to be analyzed>
/MEAN.
CSTABULATE
/PLAN FILE="< file name >"
/TABLES VARIABLES =<name of variable to be analyzed>
/CELLS TABLEPCT.
```

### *Example using SAS*

```
SAS proc surveymeans (for averages) or surveyfreq (for percentages/proportions)
PROC SURVEYMEANS;
STRATA PSTRAT;
CLUSTER PPSU;
```



```
WEIGHT WTFA_A;  
VAR <name of variable to be analyzed>;  
RUN;  
  
PROC SURVEYFREQ;  
STRATA PSTRAT;  
CLUSTER PPSU;  
WEIGHT WTFA_A;  
TABLES <name of variable to be analyzed>;  
RUN;
```

### **Example using R**

R (including the “survey” add-on package)

Note that R syntax is case-sensitive.

```
# load survey package
```

```
require(survey)
```

```
# create data frame with NHIS design information, using existing data frame of NHIS data
```

```
nhissvy <- svydesign(id=~PPSU, strata=~PSTRAT,
```

```
  nest = TRUE,
```

```
  weights=~wtfa_a,
```

```
  data=< existing data frame name>)
```

```
svymean(~<name of variable to be analyzed>,design=nhissvy)
```

Note that svymean will produce proportions for “factor variables.” For details, consult the R documentation.

### **Degrees of Freedom**

The number of degrees of freedom is used to determine the t-statistic, its associated percentage points, p-values, standard error, and confidence intervals. A rule of thumb to calculate the number of degrees of freedom to associate with a standard error is the quantity (number of PSUs - number of strata). Typically, this rule is applied to a design with at least two PSUs per stratum and when the variance components by stratum are roughly the same magnitude. This rule of thumb is not directly applicable to the NHIS design. The applicability of this rule of thumb depends upon the variable of interest and its interaction with the design structure (for additional information, see Chapter 5 of Korn and Graubard, 1999). As the number of degrees of freedom becomes large, the distribution of the t-statistic approaches the standard normal distribution. For example, with 120 degrees of freedom, the 97.5 percentage point of the t distribution is 1.980, while the 97.5 percentage point of the standard normal distribution is 1.960. If a variable of interest is distributed across most of the NHIS address clusters, a normal distribution assumption may be adequate for analysis since the number of degrees of freedom would be large. The user should consult a mathematical statistician for further discussion.

### **Variance Estimation for Subsetted Data Analysis**

Frequently, analyses using NHIS data are restricted to specific population subgroups (e.g., persons aged 65 and older). NCHS recommends that subpopulation analyses be carried out using the full data file and the SUBPOPN statement in SUDAAN, or an equivalent procedure with another complex design variance estimation software package.

Some users delete all records outside of the domain of interest (e.g., persons aged less than 65 years) in order to work with smaller data files and run computer jobs more quickly. This procedure of keeping only selected records (and list-wise deleting other records) is called subsetting the data. With a subsetted dataset that is appropriately weighted, correct point estimates (e.g., estimates of population subgroup means) can be produced. However, in general, software packages that correctly analyze complex survey data cannot compute accurate standard errors for subsetted data. When complex survey data are subsetted, often the sample design structure available to the software is incomplete; subsetting data deletes important design information needed for variance estimation.

### **Example: Variance Estimation for Subsetted Data Analysis**

The following code is to subset the second category for each variable RACEALLP\_A and SEX\_A, which happens to be the value “2” in both cases in this example. These are Sample Adult variables for race and sex where RACEALLP\_A=2 is Black/African American only and SEX\_A=2 is female.

Below are examples for subsetting NHIS data using SUDAAN, Stata, SPSS, SAS, and R software packages for illustrative purposes.

#### ***Example using SUDAAN***

SUDAAN has a SUBPOPN statement that allows the targeting of a subpopulation while using the full (unsubsetted) data file containing the design information for the entire sample.

#### **Strategy 1 (recommended)**

Use the SUBPOPN statement with the SUDAAN method described above for the full Sample Adult dataset:

```
PROC ...DESIGN = WR;  
NEST PSTRAT PPSU;  
WEIGHT WTFA _SA;  
SUBGROUP (variable names);  
LEVELS ... ;  
SUBPOPN RACEALLP_A=2 & SEX_A=2 / NAME="Analysis of African American women;"
```

Using the full dataset with the SUBPOPN statement in this example would constrain this analysis to African American women only (RACEALLP\_A = 2 for black and SEX\_A = 2 for female). Use of the SUBPOPN statement is equivalent to subsetting the dataset, except that any resulting variance estimates are based on the full design structure for the complete dataset.

#### **Strategy 2 (not recommended, except when Strategy 1 is infeasible)**

Use the MISSUNIT option on the NEST statement with the method described above for subsetted data:  
NEST PSTRAT PPSU / MISSUNIT;

In a WR design, when some PSUs are removed from the database through the listwise deletion of records outside the population of interest, leaving only one PSU in one or more strata, the MISSUNIT option in SUDAAN “fixes” the estimation to avoid errors due to the presence of strata with only one PSU. In the special case of a WR design with exactly two PSUs per stratum, using the MISSUNIT option with subsetted data gives the same variance estimate as using Strategy 1. However, except for this special case, there is no guarantee that the

variance estimates obtained by this method are equivalent to those obtained using Strategy 1. Other calculations, such as those for design effects, degrees of freedom, standardization, etc., may need to be carried out differently.

### ***Example using STATA***

Stata svy

Add SUBPOP to the SVY statement, e.g.:

```
svy, subpop( raceallp_a==2 & sex_a==2 ): mean <name of variable to be analyzed>
```

### ***Example using SPSS***

SPSS csdescriptives or cstabulate

One must first define an indicator variable, e.g.:

```
DO IF (RACEALLP_A EQ 2 AND SEX_A EQ 2).
```

```
  COMPUTE SUBGRP=1.
```

```
ELSE.
```

```
  COMPUTE SUBGRP=0.
```

```
END IF.
```

And then refer to the indicator variable in csdescriptives or cstabulate, e.g.:

```
CSDESCRIPTIVES (or CSTABULATE)
```

```
/SUBPOP TABLE=SUBGRP
```

It is very important that the indicator variable be defined for all data records. Otherwise, an invalid result can occur.

### ***Example using SAS***

SAS proc surveymeans or surveyfreq

One must first define an indicator variable, e.g.:

```
IF RACEALLP_A=2 & SEX_A=2 THEN SUBGRP=1;
```

```
ELSE SUBGRP=0;
```

And then refer to the indicator variable in proc surveymeans using the DOMAIN statement, e.g.:

```
PROC SURVEYMEANS;
```

```
DOMAIN SUBGRP;
```

Proc surveyfreq does not have a DOMAIN statement. Instead, include the indicator variable in the TABLES specification:

```
PROC SURVEYFREQ;
```

```
TABLES SUBGRP*<name of variable to be analyzed>;
```

This will produce tables for all values of the SUBGRP variable. As with SPSS, it is very important that the indicator variable is defined for all data records. Otherwise an invalid result can occur.

### Example using R

R (including the “survey” add-on package)

After applying the `svydesign` function to a data frame that contains the entire NHIS sample file being analyzed, specify the criteria that define the subgroup of interest in the `subset` function and apply the function to the R “object” created by the `svydesign` function to create a new R object. Note that the syntax that follows specifies the subgroup of interest without using an equality test.

```
# subset for raceallp_a=2 & sex_a=2 without using equal signs
subgrp <- subset(nhissvy, raceallp_a>1 & raceallp_a<3 & sex_a>1)
svymean(~<name of variable to be analyzed>, design=subgrp)
```

Note that users may want to recode variables such that missing values (which have numeric codes greater than 1) are not treated as real values. For example, `sex>1` would include missing codes 7, 8 and 9 (don’t know, refused, not ascertained respectively).

### Analysis of Subgroups

Analyses of large NHIS subgroups usually produce reliable estimates, but analyses of small subgroups may yield unreliable estimates, as indicated by their larger variances. The analyst should pay attention to the coefficient of variation (relative standard error) for estimates of means, proportions, and totals. In addition, small sample sizes, or small numbers of primary sampling units containing targeted data, may be an indication of estimates lacking precision.

## III. Merging and Appending Files

### Merging Files

NHIS data files can be merged within a survey year. The purpose of merging data within the same data year is to either incorporate variables from different data files when persons are common to both files (i.e. merging by common factor). An example of the need to merge a data file is if the user is interested in merging the Paradata file with the Sample Adult file (or the Sample Child file) to explore the associations between a wide range of methodological measures and survey data. Another example is if user is interested in merging the Sample Adult file with the Imputed Income for the Sample Adult (or the Sample Child file with the Imputed Income file for the Sample Child) to include both observed and imputed family income information in the analyses.

To merge 2019 files, use variable `HHX` as the unique identifier between the two files. Separate code is provided below to illustrate merging the Sample adult file with the Paradata file, and to merge the Sample Adult file with the Sample Adult Imputed Income file and apply an imputation variable to the analyses for the appropriate calculation of standard error of the imputed variable.

### Example: Merging survey data and Paradata using SAS

```
PROC SORT DATA=ADULT19;
BY HHX;
RUN;
```

```
PROC SORT DATA=PARADATA19;  
BY HHX;  
RUN;
```

```
/* creates a new file with Sample Adult and Paradata variables for each household */  
DATA ADULT19_PLUS_PARA;  
MERGE ADULT19 PARADATA19;  
BY HHX;  
RUN;
```

### *Example: Merging survey data and Paradata using STATA*

```
cd c:\nhis2019\  
  
use adult19  
sort hhx  
save adult19, replace  
  
use paradata19  
sort hhx  
save paradata19, replace  
  
merge 1:1 hhx using adult19  
save adultpara19 /* creates a new file with Sample Adult and Paradata variables for each household */
```

### *Using Imputed Income Data Files*

The following code is for merging the Sample Adult data file and the Sample Adult imputed income file and conduct an analytic procedure (i.e. logistic regression using survey data). The variables used in this example were recoded as illustrated in Table 2.

The analytic example, using the 2019 Sample Adult data file and the Sample Adult imputed income file, will examine the effect of the variables RATCAT\_A (the ratio of family income to the poverty threshold [*imputed*]), and WELLCHK (last doctor's visit was a wellness visit, recoded from WELLNESS\_A) on HEALTH (having good-to-excellent health, recoded from PHSTAT\_A).

Table 2. Variables in the imputed income example.

Survey question	Original variable name	Original values	Recoded variable name	Recoded values
Randomly assigned household number unique to household	HHX	Range of unique of values	(not recoded)	(not recoded)
Would you say your health in general is excellent, very good, good, fair, or poor?	PHSTAT_A	1. Excellent 2. Very good 3. Good 4. Fair 5. Poor 7. Refused 8. Not ascertained 9. Don't know	HEALTH	1. Excellent, very good or good 0. Fair or poor . (missing)
Was the last doctor's visit a wellness visit, physical, or general purpose check-up?	WELLNESS_A	1. Yes 2. No 7. Refused 8. Not ascertained 9. Don't know	WELLCHK	1. Yes 2. No . (missing)
Ratio of income to poverty threshold	RATCAT_A	1. 0.00 - < 0.50 2. 0.50 - < 0.75 3. 0.75 - < 1.00 4. 1.00 - < 1.25 5. 1.25 - < 1.50 6. 1.50 - < 1.75 7. 1.75 - < 2.00 8. 2.00 - < 2.50 9. 2.50 - < 3.00 10. 3.00 - < 3.50 11. 3.50 - < 4.00 12. 4.00 - < 4.50 13. 4.50 - < 5.00 14. >= 5.00	(not recoded)	(not recoded)

### Example: Imputed income analysis using SAS

In SAS, analysis of multiple imputed data is conducted in two stages:

1. Analysis: each of the M imputed datasets is analyzed separately using any method that would have been selected had there been a single complete dataset. This includes analytical procedure in SAS, such as PROC GLM, PROC MIXED, PROC LOGISTIC, PROC FREQ, etc.
  - In SAS, analysis of multiply imputed data is invoked with a "BY \_IMPUTATION\_" statement, to indicate that the same analysis is performed within each of the imputed datasets.
  - Users need to rename the NHIS imputation number identifier IMPNUM to \_IMPUTATION.

2. Pooling: analysis results from M imputed datasets obtained from step 1 are combined into one overall result. This step can be carried out using SAS PROC MIANALYZE.

Import data files into SAS. See SAS input statements provided on the 2019 NHIS Data Release page. This example uses the libname 'NHIS.'

```
/*The sample code below illustrates renaming the IMPNUM variable to _IMPUTATION_ for analyses in a new SAS dataset. */
```

```
DATA IMPINC;  
SET NHIS.ADULTINC19;  
RENAME IMPNUM= _IMPUTATION_; *SAS identifies imputed datasets by imputation_;  
RUN;
```

```
/* Next, merge the Sample Adult file and Sample adult imputed Income file. Data files must be sorted by the common ID before they can be merged*/
```

```
PROC SORT DATA= IMPINC;  
BY HHX;  
RUN;
```

```
PROC SORT DATA=NHIS.ADULT19 OUT=ADULT19;  
BY HHX;  
RUN;
```

```
DATA NHIS19;  
MERGE ADULT19 (IN=A) IMPINC; *Merging the imputed income and the main dataset;  
BY HHX;  
IF A;  
RUN;
```

```
/* Sort the new dataset by imputation prior to analysis. Otherwise, your analyses will only show the first category of each variable for imputed analyses */
```

```
PROC SORT DATA= NHIS19;  
BY _IMPUTATION_;  
RUN;
```

```
/*The survey analytic procedure (PROC SURVEYLOGISTIC) is used to account for the complex sampling design of NHIS. */
```

```
PROC SURVEYLOGISTIC DATA=NHIS19;  
STRATUM PSTRAT;  
CLUSTER PPSU;  
WEIGHT WTFA_A;  
CLASS RATCAT_A WELLCHK (DESC);  
MODEL HEALTH(EVENT='1') = RATCAT_A WELLCHK;  
ODS OUTPUT PARAMETERESTIMATES=lgsparms ODDSRATIOS=lgsodds;  
BY _IMPUTATION_;
```

```
RUN;
```

```
/*The ODS datasets from the code above will contain a set of estimates for each imputed dataset identified by the variable _imputation_ included in each of them.
```

```
The MIANALYZE procedure combines the results of the analyses of imputed data and generates valid statistical inferences.*/
```

```
PROC MIANALYZE PARMs(CLASSVAR=CLASSVAL)=lgsparms;  
  CLASS RATCAT_A WELLCHK;  
  MODELEFFECTS RATCAT_A WELLCHK;  
  ODS OUTPUT PARAMETERESTIMATES=mian_lgsparms; *Combines the results of previous analyses;  
RUN;
```

```
/*The mian_lgsparms output shows parameters from the pooled imputed datasets.*/
```

### *Example: Imputed income analysis using SAS-callable SUDAAN*

SUDAAN reads in separate imputed datasets. To conduct analyses in SAS-callable SUDAAN, the following steps are taken:

1. Separate the multiply imputed NHIS.ADULTINC19 SAS data set into 10 individual imputed income datasets impinc1-impinc10.
2. Merge each imputed dataset with the main NHIS data file.

Import data files into SAS. See SAS input statements provided on the 2019 NHIS Data Release page. This example uses the libname 'NHIS.'

```
PROC SORT DATA=nhis.ADULT19; *Sorting by HHX;  
BY HHX;  
RUN;
```

```
/* The following macro creates 10 separate imputation datasets from the NHIS.ADULTINC19 multiply imputed file. It then sorts them by the merge variable HHX and merges each with the NHIS.ADULT19 dataset. Ultimately 10 separate datasets are created for imputed analyses in SUDAAN */
```

```
%MACRO SEPARATE;  
%DO I= 1 %TO 10; *Instructs SAS to do the procedure for 10 iterations;
```

```
DATA IMPINC&I; *CREATING 10 SEPARATE IMPUTED DATASETS;  
SET NHIS.ADULTINC19;  
WHERE IMPNUM= &I;  
RUN;
```

```
PROC SORT DATA= IMPINC&I; *Sorting the 10 imputed datasets by HHX;  
BY HHX;  
RUN;
```



```
DATA NHIS19_&I; *Creating 10 separate analytic datasets;
MERGE NHIS.ADULT19 IMPINC&I;
BY HHX;
```

```
/* SUDAAN requires analytic datasets be sorted by the design/nest variables.
These variables are PSTRAT and PPSU on the NHIS. */
```

```
PROC SORT DATA= NHIS19_&I; *SORTING THE IMPUTED DATASETS BY DESIGN VARIABLES;
```

```
BY PSTRAT PPSU;
RUN;
```

```
%END;
%MEND;
%SEPARATE;
```

```
/*In SUDAAN, the option MI_COUNT indicates use of multiple imputed datasets.
The associated numeral indicates the number of imputed datasets to be used in the analysis. Note that the data
name used after the DATA= is the name of the first imputed dataset. This dataset name ends with the number 1.
*/
```

```
PROC RLOGIST DATA = NHIS19_1 FILETYPE=SAS DESIGN=WR MI_COUNT=10 ;
NEST PSTRAT PPSU / MISSUNIT;
WEIGHT WTFA_A;
```

```
SUBGROUP WELLCHK ;
LEVELS 2 ;
```

```
REFLEVEL WELLCHK=1 ;
```

```
MODEL HEALTH= RATCAT_A WELLCHK ;
EFFECTS WELLCHK= (2 -1)/EXP NAME="WELLCHK Yes vs No" ;
```

```
PRINT / betas=default risk=default tests=default expcntrst=default
t_betafmt=f7.2 waldfmt=f8.2 dffmt=f10.0 orfmt=f5.2 loworfmt=f5.2
uporfmt=f5.2 exp_cntrstfmt=f13.2 low_cntrstfmt=f5.2 up_cntrstfmt=f5.2;
```

```
SETENV COLWIDTH=15 DECWIDTH=4 LABWIDTH=25 COLSPCE=1 TOPMGN=0;
```

```
RLABEL HEALTH="In good-excellent health";
RTITLE "Using LOGISTIC to Model good-excellent health";
RUN;
```

### *Example: Imputed income analysis using STATA*

Import data files into Stata. See the sample Stata *.do* statements provided on the 2019 NHIS Data Release page.

```
cd c:\nhis2019\
```

```
use adult19
sort hhx
save nhis19
```

```
use adultinc19
sort hhx
save impinc
```

```
use nhis19
merge 1:m hhx using "impinc"
append using "nhis19"
```

```
// Rename the NHIS imputation number identifier impnum to _mi_m
```

```
replace impnum=0 if impnum==.
save nhis19_mi, replace
```

```
//set data to mi svyset
```

```
mi import flong, m(impnum) id(hhx)
```

```
mi svyset [pweight=WTF_A], strat(pstrat) psu(ppsu) singleunit(centered)
save nhis19_mi_dat
```

```
//mi describe will list the registration status of the variables. mi varying will report the varying and super-varying variables. Verify that all varying variables are registered as imputed or passive.
```

```
mi describe
mi varying
```

```
//logistic regression
mi estimate: svy: logistic health RATCAT_A wellchk
//odds ratios
mi estimate, or: svy: logistic health RATCAT_A wellchk
```

## Appending files

The 2019 Sample Adult and Sample Child can be appended to one another (i.e. add observations from different persons) to facilitate the analysis of measures that are common to both adults and children. An example of the need to combine observations or concatenate the Sample Adult and Sample Child files would be if the user is interested in generating an estimate of the U.S. civilian noninstitutional population of both children and adults or any subset of age ranges that includes both children and adults (e.g. ages 0 to 64 years).

To do so, data from the Sample Adult file and the Sample Child file should have comparable measures available. Since the names of the sample adult variables end in “\_A” and the names of the sample child variables end in “\_C”, comparable measures should be renamed to a common variable name. The Sample Adult and Sample Child weights should also be renamed to have the same variable name. The variance estimation variables have

the same name for both the Sample Adult and Sample Child files, and no additional recoding and renaming is needed.

The following code illustrates the concatenation of the 2019 Sample Adult and Sample Child files for the purpose of generating an estimate of the U.S. civilian noninstitutional population who are uninsured, by age. The example code illustrates keeping variables of interest for the analysis and recoding them in order to generate the estimate of interest. The code does not include analytic procedures for generating the estimate. The variables used in this example are illustrated in Table 3.

Table 3. Variables in example concatenating the Sample Adult file and the Sample Child

Variable description	Variable name in the Sample Adult file	Variable name in the Sample Child file	Variable name in new file
Indicates person is the sample adult; Indicates person is the Sample Child	HHSTAT_A	HHSTAT_C	HHSTAT_A and HHSTAT_C
Pseudo-stratum for public-use file variance estimation	PSTRAT	PSTRAT	PSTRAT
Pseudo-PSU for public-use file variance estimation	PPSU	PPSU	PPSU
Weight - Final Annual	WTFA_A	WTFA_C	WTFA_NEW
Coverage status as used in Health United States	NOTCOV_A	NOTCOV_C	NOTCOV
Age of sample adult (top coded); Age of sample child	AGEP_A	AGEP_C	AGE

### *Example: Concatenating files using SAS*

Create a new file with all Sample Adult and all Sample Child records. Keep variance and sample weights, common variables in both files and new recodes combining key variables.

```
DATA PERSON19;
SET ADULT19 CHILD19;
KEEP
PSTRAT PPSU WTFA_A WTFA_C HHSTAT_A HHSTAT_C
NOTCOV_A AGEP_A
NOTCOV_C AGEP_C
WTFA_NEW
NOTCOV
AGE;

/* recodes*/
IF HHSTAT_A='1' THEN DO;
WTFA_NEW=WTFA_A;
NOTCOV=NOTCOV_A;
AGE = AGEP_A;
END;
```

```
IF HHSTAT_C='1' THEN DO;
WTFA_NEW=WTFA_C;
NOTCOV=NOTCOV_C;
AGE = AGEP_C;
END;
RUN;
```

### *Example: Concatenating files using STATA*

Create new separate files for the sample adult and sample child with the variables for variance, sample weight, and common variables of interest. Combine files and recode combining key variables.

```
use child19
keep NOTCOV_C HHSTAT_C ppsu pstrat WTFA_C AGEP_C
save childvars

use adult19
keep NOTCOV_A HHSTAT_A ppsu pstrat WTFA_A AGEP_A
save adultvars

append using childvars

//Recodes
gen notcov=.
replace notcov=1 if NOTCOV_C==1 | NOTCOV_A==1
replace notcov=2 if NOTCOV_C==2 | NOTCOV_A==2

gen age=.
replace age=AGEP_C if HHSTAT_C==1
replace age=AGEP_A if HHSTAT_A==1

gen WTFA_new=.
replace WTFA_new= WTFA_C if HHSTAT_C==1
replace WTFA_new= WTFA_A if HHSTAT_A==1
save vars_child_adult
```

## IV. Section Acronyms and Questionnaire Order

### Acronym Definitions

Table 4. Acronym definition of sections in the questionnaire and codebook documents: 2019 NHIS

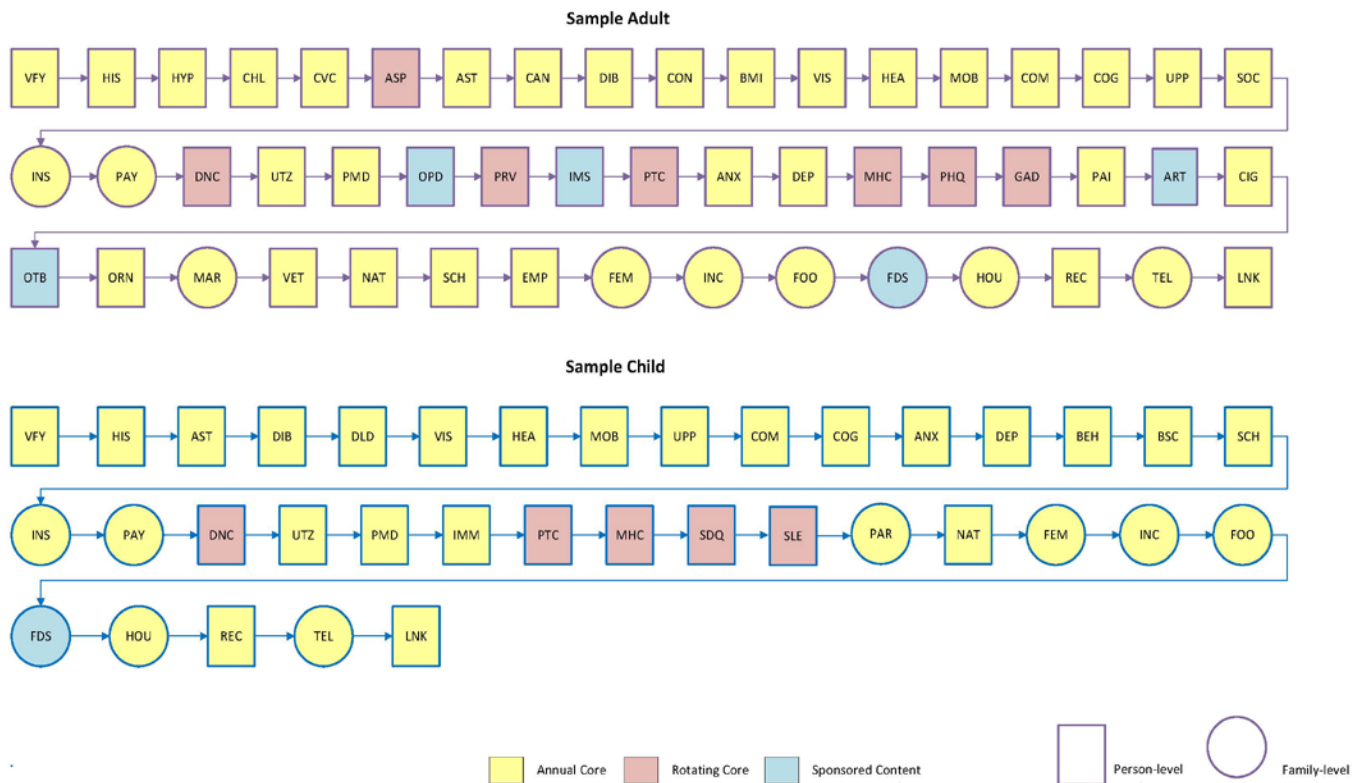
Section Acronym	Description
ANX	Anxiety
ART	Arthritis
ASP	Aspirin
AST	Asthma
BEH	Behavior
BMI	Current Pregnant, Height, Weight
BPS	Baby Pediatric Symptom Checklist
CAN	Cancer
CHL	Cholesterol
CIG	Cigarettes and E-cigarettes
COG	Cognition
COM	Communication
CON	Other Conditions
CVC	Cardiovascular Conditions
DEP	Depression
DIB	Diabetes
DLD	Developmental and Learning Disabilities
DNC	Dental Care
EMP	Employment of Sample Adult
FDS	Food Security
FEM	Employment of Family Members
FOO	Food Related Programs
GAD	Anxiety
HEA	Hearing
HHC	Household Composition
HIS	Health Status
HOU	Housing
HYP	Hypertension
IMM	Immunization (Sample Child)
IMS	Immunization (Sample Adult)
INC	Family Income
INS	Health Insurance
LNK	Linkage
MAR	Marital Status
MHC	Mental Health

Table 4. Acronym definition of sections in the questionnaire and codebook documents: 2019 NHIS, continued.

<b>Section Acronym</b>	<b>Description</b>
MOB	Mobility
NAT	Nativity
OPD	Prescription Opioid Use
ORN	Sexual Orientation
OTB	Other Tobacco
PAI	Chronic Pain
PAR	Parent Demographics
PAY	Difficulty Paying for Insurance
PHQ	Depression
PMD	Prescription Medication
PRV	Preventive Screening
PTC	Physical and Other Therapeutic Care
REC	Adult's Full Name
SCH	Schooling
SDQ	Strengths and Difficulties Questionnaire
SLE	Child Stressful Life Events
SOC	Social Functioning
TEL	Telephone Use
UPP	Self-Care and Upper Body
UTZ	Utilization
VET	Veterans Status
VFY	Verification and Demographic Details
VIS	Vision

### Order of 2019 Questionnaire Sections

Figure 3. Questionnaire section order of Sample Adult and Sample Child Modules, by type of question: 2019 NHIS



## Sample Adult's Health

### I. Health Status and Conditions

#### Annual Core

Several sections throughout the Sample Adult module measure the health of U.S. adults. Sample adults were asked to self-report their height and weight, self-perceived health status, current pregnancy status for females aged 18-49 years, and whether a doctor or other health care professional had told them that they had series of selected conditions. Estimates derived from questions that ask about specific health conditions diagnosed by a doctor or health care professional may underestimate the true burden of these conditions in the population due to the undiagnosed status of the condition during its detectable pre-clinical and clinical phase and from reporting bias. For a list of health conditions measured in the annual core, by questionnaire section and reference periods of its respective questions, see Table 5. Sample adults were also asked about the intake of medication to treat diabetes, hypertension and high cholesterol, and visits to an emergency room due to asthma, if ever diagnosed with these respective conditions. Age of diagnosis was collected from sample adults ever diagnosed with diabetes and cancers.

Table 5. Health conditions about the sample adult measured in the annual core, by questionnaire section and reference periods of its respective questions.

Topic	Section	Reference period in available questions
Angina pectoris	CVC	Ever
Anxiety disorder	CON	Ever
Arthritis, rheumatoid arthritis, gout, lupus, or fibromyalgia	CON	Ever
Asthma	AST	Ever; Past 12 months; Current
Cancer and cancer kind	CAN	Ever
Chronic Obstructive Pulmonary Disease	CON	Ever
Coronary heart disease	CVC	Ever
Dementia, including Alzheimer's disease	CON	Ever
Depression	CON	Ever
Diabetes	DIB	Ever
Gestational diabetes	DIB	Ever
Health status	HIS	Current
Height and weight	BMI	Current; If pregnant, before pregnancy
Hyperlipidemia	CHL	Ever; Past 12 months
Hypertension	HYP	Ever; Past 12 months
Myocardial infarction	CVC	Ever
Pre-diabetes	DIB	Ever
Pregnancy status	BMI	Current
Stroke	CVC	Ever



---

## Sponsored Content

NCCDPHP and NCIRD sponsored additional information on arthritis in the Sample Adult ART section. Sample adults who reported having been told by a doctor or health professional that they had arthritis, rheumatoid arthritis, gout, lupus, or fibromyalgia were asked questions about whether arthritis limited their activities, if they were advised by a doctor or health professional to lose weight and exercise, and joint pain experienced in the past 30 days.

NCCDPHP and NIDDK sponsored additional information on insulin use in the Sample Adult DIB section. Sample adults who reported having been told by a doctor or health professional that they had diabetes and are now taking insulin were asked about when insulin began and whether it was discontinued after receiving a diabetes diagnosis.

## Major Recodes

**Age of cancer diagnosis.** The age when first told by a doctor or health professional that the sample adult had a specific kind of cancer is available for 27 of the 29 different kinds of cancers and 'other' kind of cancer collected in NHIS. To protect confidentiality, sample adults mentioning kidney or testicular cancer were suppressed along with the age of diagnosis for these specific cancers. The age of cancer diagnosis for those mentioning kidney and testicular cancer are included in the recode for age of diagnosis for 'other' kind of cancer. The age of cancer diagnosis was also calculated for combined cancer types colorectal (which combines colon and rectal cancer) and for head and neck cancers (which combines larynx-tracheal, mouth/tongue/lip, and throat (pharyngeal) cancers). The youngest age provided for the respective cancer type was assigned for sample adults with combined cancers, those who identified the same kind cancer as the second or third kind of cancer, or mentioned 'other kind of cancer' more than once. Sample adults who reported 'don't know' or 'refused' when asked about the type of cancer but provided an age of diagnoses for the unidentified cancer are included in the variable for age of cancer diagnosis for 'other kind of cancer.' Variables for age of cancer diagnosis are top-coded at age 85 for all cancers, and bottom-coded at age 18 for cancers of the breast, cervix, prostate, and skin (nonmelanoma, melanoma and unknown type).

**Body Mass Index (BMI).** A categorical measure of BMI was created using unrestricted height and weight values which contain the greater range of height and weight values than are available on the public use file. BMI was calculated using the formula:  $BMI = \text{kilograms} / \text{meters}^2$  where 1 kilogram = 2.20462 pounds and 1 meter = 39.37008 inches. The categorical measure of BMI was classified as follows: underweight is  $BMI < 18.5$ ; healthy weight is  $BMI 18.5$  to  $< 25$ ; overweight is  $BMI \geq 25$  to  $< 30$ ; and obese is  $BMI \geq 30$ . Sample adults who answered don't know or refused for height or weight, reported values for either height or weight that were outside the limits for public data release, or for whom height or weight values for public release were set to 96 and 996 due to missing sex information were coded as 99 in the categorical BMI measure.

**Cancer types.** Sample adults could name up to three kinds of cancer. Twenty-nine kinds of cancer could be identified in NHIS, with the option for 'other' not listed. Responses from the up to three kinds of cancers were recoded during editing into "mentioned"/ "not-mentioned" variables for each cancer type. Due to confidentiality concerns, recodes that identified whether kidney cancer or testicular cancer were mentioned were suppressed, and sample adults reporting these cancers were included in the recode 'other cancer type' mentioned. Recodes were also created that combine specific kinds of cancers available in the public use file. Sample adults reporting colon or rectal cancer were combined into a separate variable indicating whether colorectal cancer was "mentioned"/ "not-mentioned." Similarly, Sample adults reporting larynx-tracheal,

mouth/tongue/lip, or throat (pharyngeal) cancer were recoded into a separate variable indicating whether a head and neck cancer was mentioned/not mentioned. Sample adults who reported having had cancer but reported 'don't know' or 'refused' when asked about the type of cancer are recoded as "don't know" or "refused" in these recodes, respectively. Respondents who reported a combination of "don't know" and "refused" to type of cancers are recoded as "don't know" in these recodes.

**Height and Weight.** Sample adults had the option to report height and weight using the U.S. customary system (pounds: feet and inches) or the metric system (kilograms: meters and centimeters). Metric responses on height and weight were converted into the U.S. Customary system using the following conversion scale: 1 meter=39.37008 inches and 1 kilogram=2.20462 pounds (lbs.). The public use height variable reflects total height in inches (e.g. 65" is 5' 5"), with height ranges 62-76 inches for men and 58-70 inches for women. The public use weight variable reflects total pounds rounded to whole integers with weight ranges 126-299 lbs. for men and 100-274 lbs. for women. Pregnant women aged 18-49 were asked to report weight before pregnancy. Sample adults who reported values outside the public use limits for either height or weight had data for both variables recoded to "96" or "996" ("Not available") on the public use data file to protect the confidentiality of those who might be identifiable by their unusual physical characteristics. In addition, due to the sex-specific height and weight limits for public data release, all sample adults for whom sex was answered as don't know or refused were also coded as 96 for height and 996 for weight. Don't know and refused responses to either height or weight were retained in the height and weight public use recodes.

**Number of cancers.** A summary recode was created that indicates the number of different kinds of cancer mentioned by the sample adult. Sample adults who were never told that they had cancer were assigned the value 0. The number of different kinds of cancers were derived from responses to the type of cancer told to have and whether they had a second, third and other (additional) types of cancer (assigned values 1-4, respectively, where 4 indicates four or more).

**Years since diabetes diagnosis.** This recode is the calculation of the sample adult's current age (in years) minus the age (in years) when first told to have diabetes. The highest number of years since diabetes diagnosis available in the public data release is 85 years, and years since diagnosis for sample adults aged 85 or older is collapsed into one or more years and coded 96.

## II. Functioning and Disability

### Annual Core

#### *Functioning and Disability*

The questions on functioning and disability found in sections VIS, HEA, MOB, COM, COG, UPP, ANX, DEP, PAI, and FGE of the sample adult module are part of sets of international standard measures developed, tested and endorsed by the Washington Group on Disability Statistics (WG). The WG is a city group established in 2001 under the United Nations Statistical Commission to address the need for population-based measures of disability by promoting and coordinating international cooperation in the area of health statistics focusing on disability data collection tools suitable for censuses and national surveys. The major objective is to provide necessary information on disability that is comparable throughout the world by identifying individuals with functional limitations in basic actions, regardless of nationality or culture. The questions reflect advances in the conceptualization of disability and use the World Health Organization's International Classification of

Functioning, Disability, and Health (ICF) as a conceptual framework. The intended use of these questions is to describe the functional status of adults and, when used with other questions on the survey, to evaluate whether adults with functional limitations have achieved similar levels of participation and inclusion as adults without functional limitations. These questions do not capture all aspects of difficulty in functioning, but rather focus on domains of functioning that are likely to identify the majority of adults at risk of participation restrictions in an unaccommodating environment.

Table 6. Functional limitations measured in the Sample Adult module, by section and Washington Group (WG) Set on Functioning

Topic	Section	WG-SS	WG-SS Enhanced	WG-ES
Wear glasses or contacts	VIS			x
Have difficulty seeing	VIS	x	x	x
Use a hearing aid	HEA			x
How often use hearing aid	HEA			x
Have difficulty hearing	HEA	x	x	x
Difficulty walking or climbing steps	MOB	x	x	x
Use equipment or receive help for getting around	MOB			x
Type of equipment	MOB			x
Difficulty walking 100 yards	MOB			x
Difficulty walking a third of mile	MOB			x
Difficulty walking up or down 12 steps	MOB			x
Difficulty communicating	COM	x	x	x
Difficulty remembering or concentrating	COG	x	x	x
Difficulty remembering, concentrating, or both	COG			x
How often have difficulty remembering	COG			x
Difficulty remembering few things, a lot, everything	COG			x
Difficulty with self-care	UPP	x	x	x
Difficulty raising a 2-liter bottle from waist to eye level	UPP		x	x
Difficulty using hands and fingers	UPP		x	x
How often feel worried, nervous, or anxious	ANX		x	x
Take prescription medication for these feelings	ANX			x
Level of feelings of worried, nervous, anxious	ANX		x	x
How often feel depressed	DEP		x	x
Take prescription medication for these feelings	DEP			x
Level of feelings of depressed	DEP		x	x
How often have pain	PAI			x
How much pain	PAI			x
How often felt very tired or exhausted*	FGE			x
How long tired or exhausted feelings last*	FGE			x
Level of tiredness or exhaustion*	FGE			x

\*Questions about fatigue are not included in the 2019 NHIS and are part of the rotating core content starting in 2020.

NOTES: WG-SS is WG Short Set on Functioning, WG-SS Enhanced is WG Short Set on Functioning – Enhanced, and WG-ES is WG Extended Set on Functioning.

The questions included for sample adults are from the WG Extended Set on Functioning (WG-ES). Two additional question sets developed by the WG are subsets of the WG-ES set – the WG Short Set on Functioning (WG-SS) and the WG Short Set on Functioning – Enhanced (WG-SS Enhanced), comprised of 6 and 12 questions, respectively. For a list of questions included in each set and their respective questionnaire sections, see Table 6.

Questions ask about the sample adults' level of difficulty (no difficulty, some difficulty, a lot of difficulty, or cannot do at all) in basic domains of functioning including seeing, hearing, mobility, communication, cognition, self-care, and upper body functioning and about the frequency and intensity of experiencing anxiety and depression, pain and fatigue. In addition to questions about level of difficulty, several of the domains have questions to provide information on the use of accommodations. The WG questions on pain are described below in the pain and management section. Questions on fatigue are part of the rotating core starting in 2020. More information on the Washington Group and the question sets may be obtained by request to the WG Secretariat at [WG\\_Secretariat@cdc.gov](mailto:WG_Secretariat@cdc.gov) or found on the WG website at: <http://www.washingtongroup-disability.com/>.

The WG questions can be analyzed separately, by domain, or combined across domains. A disability status indicator is available to data users (See Major Recodes below) that identifies sample adults who are at greater risk than the general population for experiencing restrictions in participation because of difficulties doing certain universal, basic actions according to the internationally agreed upon definition. The recode identifies adults who “have a lot of difficulty” or “cannot do at all” to at least one of the six domains included in the WG Short Set on Functioning. Use of the functioning and disability data should be tailored to the needs of the analysis. Other disability indicators can be created from the WG Extended Set on Functioning and the WG Short Set Enhanced. Consult the WG website for guidance on the creation of these indicators. Changing the threshold for either the number of domains the respondent identifies having difficulty with or the degree of difficulty can create different identifiers that will capture different populations. For example, a recode that includes respondents who have “some difficulty” with any of the domains will capture a greater proportion of the population than a recode limited to include only those who report they “cannot do at all” to any of the domains and the functional abilities of the larger group will be much more heterogeneous than that of the smaller group. Analytic guidelines, including recommended disability identifiers, written for each of the WG questions sets may be obtained from the WG website: <http://www.washingtongroup-disability.com/washington-group-question-sets/short-set-of-disability-questions/>.

## Participation

The SOC section in the Sample Adult module contains three questions about difficulty with participation in everyday life activities that are not part of the WG questions on functioning. These questions directly ask about participation by determining whether because of a physical, mental, or emotional condition the sample adult is limited in the kind or amount of work the respondent can perform, and the degree of difficulty the sample adult has doing errands alone, or participating in social activities. Analysts can analyze each question separately or can combine the questions into an indicator to meet the needs of their analysis.

## Major Recodes

**Disability status composite indicator, age 18 and older.** An indicator of disability status based on the WG Short Set on Functioning that identifies sample adults who are at greater risk than the general population for experiencing restrictions in participation because of difficulties doing certain universal, basic actions. This recode classifies sample adults with disability as those reporting “a lot of difficulty” or “cannot do at all” for at least one of six domains of functioning: seeing (even if wearing glasses), hearing (even if wearing hearing aids), mobility

(walking or climbing stairs), communication (understanding or being understood by others), cognition (remembering or concentrating), and self-care (such as washing all over or dressing). The remaining sample adults, that is those who responded "some difficulty" or "no difficulty" to at least one domain (and did not report "a lot of difficulty" or "cannot do at all" for any of the six domains of functioning) are classified as without disability. Those responding "don't know" or "refused" to all six questions are excluded.

### III. Pain and Pain Management

#### Rotating Core

Beginning in 2019 and rotating in every other year are questions about pain frequency, amount of pain, pain interference (how much pain impacts daily life and family members), and specific pain locations. These questions are an extension of the work of National Pain Strategy, an Interagency Pain Research Coordinating Committee at the National Institutes of Health (<https://www.iprcc.nih.gov/National-Pain-Strategy/Objectives-Updates>). The questions in the PAI section are designed to develop consistent population-based estimates of chronic pain and high impact chronic pain (defined as chronic pain that frequently limits life or work activities), as well as amount of pain experienced for common pain locations.

The rotating core PAI questions in the sample adult module ask about frequency of pain in the past 3 months, the amount of pain the last time they experienced pain, how often pain limited their life or work activities in the past 3 months, and how often their pain had an effect on their family and significant others in the past 3 months. The pain location questions ask about the amount of pain experienced in the following locations: back; hands, arms or shoulders; hips, knees, or feet; headache or migraine; abdominal, pelvic, or genital pain; and toothache or jaw pain. The universe of the questions about pain limiting activities, affecting family or significant others, and pain location are restricted to sample adults who reported having pain at least some days. Analysts can analyze the pain questions separately. Two of the pain questions, frequency of pain in the past 3 months and amount of pain experienced, are also measures of the pain domain in the WG extended set of disability identifiers and can be analyzed as part of the WG Extended Set on Functioning (see Table 6).

#### Emerging Content

As another extension of the National Pain Strategy, questions on pain management techniques appear as emerging content on the 2019 NHIS. These questions ask sample adults who reported experiencing pain some days, most days, or every day in the past 3 months about their use of various techniques to manage their pain within the past 3 months. These pain management techniques include: physical therapy, rehabilitative therapy, or occupational therapy; spinal manipulation or other forms of chiropractic care; talk therapies such as cognitive-behavioral therapy (CBT); a chronic pain self-management group; yoga or Tai Chi; massage; meditation, guided imagery, or other relaxation techniques; or other method to manage pain.

The opioid use (OPD) section covers emerging content in NHIS about the use of prescription opioid pain relievers. The purpose of these questions is to examine the relationship between prescription opioid use, pain, and health outcomes measured on the NHIS. Prescription opioid pain relievers refer to pain relievers prescribed by a doctor, dentist, or other health professionals containing opioids. Examples of opioid pain relievers include hydrocodone, Vicodin, Norco, Lortab, oxycodone, OxyContin, Percocet, and Percodan. The OPD questions ask sample adults who have taken prescribed medication in the past 12 months about use of any opioid pain reliever prescribed by a medical professional in the past 12 and 3 months. Additionally, they ask about prescription opioid use for short term or acute pain, use for long term or chronic pain, and frequency of use.

## IV. Health Care Access and Health Service Utilization

### Annual Core

Several sections throughout the sample adult module measure access to and use of health services, as well as affordability of care. For a list of measures on these topics asked in the annual core, by questionnaire section and reference period of its respective questions, see Table 7. Similar content is also available for sample children.

Table 7: Measures of health care access, service use, and affordability of care asked of the sample adult in the annual core, by questionnaire section and reference periods

Measure	Section	Reference period
<b>Immunizations</b>		
Flu vaccine	IMS	Last 12 months, Month and Year
Pneumonia vaccine	IMS	Ever
Number of pneumonia vaccines received	IMS	Ever
<b>Medical Care</b>		
Saw a doctor for medical care	UTZ	Last time interval
Medical and wellness visit combined	UTZ	-
Wellness visit	UTZ	Last time interval
Usual place to go for medical care	UTZ	Current
Kind of place for medical care	UTZ	Current
Number of urgent care visits	UTZ	Last 12 months
Number of emergency department visits	UTZ	Last 12 months
Any overnight hospitalization	UTZ	Last 12 months
Delayed medical care due to cost	UTZ	Last 12 months
Needed but did not get medical care due to cost	UTZ	Last 12 months
<b>Mental Health Care</b>		
Took medication for emotions/mental health	MHC	Last 12 months
Received therapy or counseling	MHC	Last 12 months; current
Delayed getting therapy/counseling due to cost	MHC	Last 12 months
Needed but did not get therapy/counseling due to cost	MHC	Last 12 months
<b>Prescription Medication</b>		
Took prescription medication	PMD	Last 12 months
Skipped doses to save money	PMD	Last 12 months
Took less medication to save money	PMD	Last 12 months
Delayed filling prescription to save money	PMD	Last 12 months
Needed but did not get prescription due to cost	PMD	Last 12 months
<b>Problems Paying Medical Bills</b>		
Anyone in family having problems paying medical bills*	PAY	Last 12 months
Have bills unable to pay at all*	PAY	Current
Level of worry about paying medical bills if sick/accident	PAY	-

\*These are family-level replicate questions asked once per family.

## Rotating Core

For two consecutive years every three years starting in 2019, there is additional content about the use of selected health services. These questions ask sample adults about receiving dental, eye, home, and physical or other therapeutic care, and about affordability of dental care. Table 8 lists rotating core measures of health care use, and affordability, by questionnaire section and reference period of its respective questions.

Table 8. Measures of health care access, service use, and affordability of care asked of the sample adult in the rotating core, by questionnaire section and reference periods: 2019 NHIS

Measure	Section	Reference period
<b>Dental Care</b>		
Received a dental cleaning/exam	DNC	Last time interval
Delayed dental care due to cost	DNC	Last 12 months
Needed but did not get dental care due to cost	DNC	Last 12 months
<b>Physical and Other Therapeutic/Specialist Care</b>		
Received eye exam from eye specialist	PTC	Last 12 months
Received physical/speech/rehabilitative/occupational therapy	PTC	Last 12 months
Received home care	PTC	Last 12 months

## Sponsored Content

The IMS section includes several questions sponsored by NCIRD regarding CDC recommended vaccinations and exposure to health care settings. Questions about vaccinations included additional content about flu and content about shingles, tetanus and human papillomavirus (HPV) vaccines. In specific, sample adults were asked about:

Flu vaccine (Female sample adults aged 18-49 or age unknown)

- Pregnancy status during most recent flu season
- Flu shot was received before/during/after a current or recent pregnancy

Shingles vaccine (Sample adults aged 50 and over)

- Receipt of any shingles vaccine
- What kind of shingles vaccine was received (Shingrix or Zostavax)
- When each vaccine was received
- Number of Shingrix vaccines received

Tetanus vaccine

- Receipt of a tetanus shot in the last 10 years
- Whether tetanus shot included pertussis (TDAP)
- Receipt of a TDAP shot during pregnancy

HPV vaccine (Sample adults 18-64)

- Receipt of HPV vaccine
- Age of first HPV vaccine



Flu vaccination timing in relation to pregnancy status was determined as follows:

Female sample adults between 18 and 49 years old (or whose age was not known) who reported that they were currently pregnant (asked previously in the BMI section), and who had received a flu vaccine in the past 12 months, and were interviewed from January through March or from August through December were asked: “Did you get a flu vaccination before or during your current pregnancy?” Female sample adults 18-49 years (or whose age is not known) and reported that they were not currently pregnant (or pregnancy status is not known), or those who were currently pregnant and were interviewed between April-July, were asked about pregnancy status during August-March as follows: since August 1<sup>st</sup> of last year if interviewed between January-March; from August of last year through March of current year if interviewed April-July; and since August 1 of current year if interviewed August-December. Those who reported to be currently pregnant or had a recent pregnancy during August-March, and who had received a flu vaccine in the past 12 months were asked: “Earlier you said you were pregnant sometime [since August 1<sup>st</sup>, {prior year}/from August {prior year} through March {current year}/since August 1<sup>st</sup>, {current year}] Did you get a flu vaccination before, during, or after your pregnancy?”

NCIRD sponsored a question for female sample adults between 18 and 49 years old (or whose age was not known) that asked if they had a live birth in the past 12 months.

NCIRD also sponsored two questions for sample adults aged 18 and over about exposure to health care settings that ask whether, in their work or volunteer activities, the sample adult provides direct medical care to patients, or whether they do any work or volunteer activities in a health care facility.

## Major Recodes

**Year of receipt of shingles vaccine.** Recodes were created with a lower limit of 2006 for the year of the most recent Zostavax<sup>®</sup> vaccine, and 2017 for the year of the most recent Shingrix<sup>®</sup> vaccine. Sample adults reporting years earlier than these lower limits were assigned the value 9996.

**Age of first HPV vaccine.** A recode was created for age of first HPV vaccine, with a lower limit of age 9, and responses younger than age 9 were assigned the value 96.

## V. Preventive Care and Services

### Rotating Core

Rotating every other year starting in 2019 are questions about aspirin use, history of colorectal cancer screenings, and clinical checks of arterial blood pressure, glucose and cholesterol. The United States Preventive Services Task Force (USPSTF) recommends initiating low-dose aspirin use for the primary prevention of cardiovascular disease (CVD) and colorectal cancer in adults aged 50 to 59 years who have an increased CVD risk, are not at increased risk for bleeding, have a life expectancy of at least 10 years, and are willing to take low-dose aspirin daily for at least 10 years (Bibbins-Domingo and USPST, 2016 ). In the ASP section, sample adults aged 40 and over were asked if a doctor or other health professional had ever told them to take a low-dose aspirin, current use, cessation of aspirin use, or if on their own they took a low-dose aspirin each day to prevent or control heart disease.

Health information asked in the PRV section of the Sample Adult module was age restricted. In separate questions, sample adults aged 18 and over were asked about their most recent blood pressure, cholesterol, and high blood sugar or diabetes check by a doctor, nurse, or other health professional. The PRV section of the



questionnaire also asked sample adults aged 40 years and older about colorectal cancer screening. There are several tests associated with colorectal cancer screening. Sample adults were asked if they ever had a colonoscopy, sigmoidoscopy, or both, a CT colonography, fecal immunochemical or FIT test and how long since their most recent exam. The USPSTF recommends periodic screening exams for colorectal cancer (Qaseem et al., 2019).

## Sponsored Content

NIH and NCCDPHP, CDC sponsored additional content about colorectal cancer screening asked in the PRV section. This included additional questions for sample adults aged 40 years and over such as the main reason for the most recent colonoscopy and if they paid none, part, or all the cost. Sample adults were also asked about the use of Cologuard for colorectal screening.

The question of when the most recent Cologuard test occurred (CGUARDWHEN\_A) was not included in the public use file due to a programming error in the questionnaire. Approximately 20% of sample adults who ever had a Cologuard test were not asked this question. Sample adults whose most recent stool or FIT test was not a Cologuard test or who reported don't know or refused to whether their last stool or FIT test was a Cologuard test were incorrectly left out of the universe.

NIH and NCCDPHP, CDC also sponsored content about breast, cervical, and prostate cancer screenings asked in the PRV section. These included questions regarding the following: Pap smear tests, HPV tests, and hysterectomy for sample adult women 18 years and over; Mammograms for sample adult women aged 30 years and over; and prostate specific antigen (PSA) tests for sample adult men aged 40 years and over. For each type of cancer screening exam, sample adults who indicated that they had the exam were asked when the most recent screening exam occurred (time interval grouping). Sample adults were also asked their reasons for having the exam.

## VI. Health-Related Behaviors

### Annual Core

Sample adults were asked about cigarette smoking and the use of electronic cigarettes in the CIG section. All adults were asked if they had smoked at least 100 cigarettes in their entire life. Those who had done so were asked whether they now smoke cigarettes every day, some days, or not at all. Every day smokers were asked about the number of cigarettes smoked every day, while some day smokers were asked the number of days that they smoked in the past 30 days and the average number of cigarettes smoked on those days. All sample adults were asked about the use of electronic cigarettes or other electronic vaping products, even one time in their entire life. The question included related terms and a description of electronic cigarettes. Those who had ever used an electronic cigarette were asked whether they now use this product every day, some days, or not at all.

### Sponsored Content

The FDA sponsored seven questions about the use of cigars, pipes and smokeless tobacco products in the OTB section. In separate questions, all sample adults were asked about whether they had ever smoked a cigar, smoked a pipe filled with tobacco, or used smokeless tobacco products. Questions included examples and

descriptions of these products. Those who said “yes” to each respective tobacco product were asked whether they now use it every day, some days or not at all. Sample adults who had ever smoked any type of cigar were also asked about the number of days they smoked cigars in the past 30 days.

For additional information about the historical context of tobacco use questions in NHIS, see <https://www.cdc.gov/nchs/nhis/tobacco.htm>

## Major Recodes

**Cigarette smoking status.** Sample adults were classified in terms of their lifetime and current cigarette smoking status. The recode includes the following categories: current everyday smoker; current some day smoker; former smoker; never smoker; smoker, current status unknown; and unknown if ever smoked. Former smoker is defined as a person who has smoked at least 100 cigarettes in their entire life and currently does not smoke at all. Never smoker is a person who has never smoked any cigarettes or has smoked less than 100 cigarettes in their entire life. Smoker, current status unknown is defined as a person who has smoked 100 cigarettes in their entire life and the question about current smoking practices was answered as don’t know or refused, or it was not ascertained. Unknown if ever smoked includes those whose response to ever having smoked at least 100 cigarettes in their entire life was answered as don’t know or refused, or it was not ascertained.

**Electronic cigarette use status.** Sample adults were classified in terms of their ever and current electronic cigarette use. The recode includes the following categories: current e-cigarette user; used e-cigarette, not current user; never e-cigarette user; e-cigarette user, current status unknown; and unknown if ever used e-cigarettes. Current e-cigarette user is defined as a person who uses electronic cigarettes everyday or somedays. Not current user is defined as a person who has ever used an electronic cigarette even one time in their entire life and who currently does not use them at all. Never user is defined as a person who has never used electronic cigarette. User, current status unknown is defined as person who has used an electronic cigarette even one time in their entire life and the question about current use was answered as don’t know or refused, or it was not ascertained. Unknown if ever used electronic cigarettes includes those whose response to ever having used an electronic cigarette in their entire life was answered as don’t know or refused, or it was not ascertained.

## VII. Mental Health

### Rotating Core

Rotating every three years starting in 2019 are two mental health scales measuring symptoms of generalized anxiety disorder (GAD) and depression.

Symptoms of generalized anxiety disorder are measured using the 7-item Generalized Anxiety Disorder scale (GAD-7) (Spitzer et al., 2006). This is a brief scale to screen for GAD symptoms in the past two weeks and assess its severity in clinical settings and the general population (Rutter and Brown 2016, Löwe et al., 2008). GAD-7 has also been found to have moderately good operating characteristics for three other anxiety disorders – panic disorder, social anxiety disorder, and post-traumatic stress disorder (Kroenke et al., 2007). The GAD-7 was developed based on the most correlated items with a 13-item scale that included 9 items from the criteria for GAD in the Diagnostic and Statistical Manual for Mental Disorders, Fourth Edition (DSM-IV) and 4 items on the basis of review of existing anxiety scales (Spitzer et al., 2006).

Symptoms of depression are measured using the 8-item Patient Health Questionnaire depression scale (PHQ-8) (Kroenke et al., 2009). This is a valid diagnostic and severity measure for current depressive disorders derived using the nine-item criteria for depressive disorders in the DSM-IV. Symptoms of depression are asked in

reference to the past two weeks. PHQ-8 is used in clinical settings and in population-based studies to screen for symptoms indicative of a possible clinically significant depression and assess severity of depressive disorders (Kroenke et al., 2009). PHQ-8 is an abbreviated version of the nine-item PHQ-9 scale (Kroenke et al., 2001) that excludes the question about thoughts of death and self-injury, an indicator of possible suicide risk.

The set of questions in the GAD-7 and PHQ-8, available in the GAD and PHQ sections of the questionnaire, ask sample adults to assess how often they have been bothered over the last 2 weeks by a set of specific symptoms. Response options to the GAD-7 and PHQ-8 questions are the same: 1) not at all, 2) several days, 3) more than half the days and 4) nearly every day. These response categories correspond to 0 to 3 points, respectively, when scoring each question. The points for each question are summed to produce a total score between 0 and 21 in GAD-7 and between 0 and 24 for PHQ-8. Total scores may be used in analysis, and pre-determined cut-off values of the total score are available for assessing severity of symptoms (Kroenke et al., 2010).

Variables classifying the total score into four categories of severity of symptoms of GAD and depression (i.e. none or minimal, mild, moderate, and severe) are available to data users (see major recodes, below). The recode for severe symptoms of depression is based on a PHQ-8 total score of 15 or greater. The PHQ-8 total score can also be classified into five categories of severity, where those scoring 15 or greater may be further categorized into moderately severe (point values 15 to 19) and severe symptoms (point values 20 to 24) (Kroenke et al., 2009). For both GAD-7 and PHQ-8, a total score of 10 or greater is indicative of a possible clinically significant condition.

Some other surveys do not include all 8 questions of the PHQ-8 or 7 questions of the GAD-7, but instead include only the first two items of each scale. These abbreviated measures are known as the PHQ-2 and GAD-2, respectively (Kroenke, et al., 2003; Kroenke et al., 2007). PHQ-2 and GAD-2 scores can be calculated from the NHIS for comparison with other surveys. The score of the two questions per scale are summed to produce a total score between 0 and 6. A score of 3 points or greater identifies possible cases for major depressive disorder based on the PHQ-2 and for generalized anxiety disorder based on the GAD-2.

## Major Recodes

**Severity of GAD symptoms.** Severity of GAD symptoms were categorized into no or minimal symptoms, mild symptoms, moderate symptoms, and severe symptoms for sample adults with responses to at least 6 of the 7 questions, using the cut-offs of the total score at <5, <10, <15, and 15 or greater. Instances when one question was refused or answered as don't know, or it was not ascertained, the average of the 6 ascertained answers (with point values 0-3) was calculated and added to the total score from the six ascertained questions. Total scores with decimal values were not rounded-up to the next whole integer for determining category of severity. Sample adults missing responses to two or more questions were categorized as having unknown severity of symptoms and coded as 9.

**Severity of depression symptoms.** Severity of depression symptoms were categorized into no or minimal symptoms, mild symptoms, moderate symptoms, and severe symptoms for sample adults with responses to at least 7 of the 8 questions, using the cut-offs of the total score at <5, <10, <15, and 15 or greater. Instances when one question was refused or answered as don't know, or it was not ascertained, the average of the 7 ascertained answers (with point values 0-3) was calculated and added to the total score from the seven ascertained questions. Total scores with decimal values were not rounded-up to the next whole integer for determining category of severity. Sample adults missing responses to two or more questions were categorized as having unknown severity of symptoms and coded as 9.

---

## Sample Child's Health

### I. Health status and conditions

#### Annual Core

The Sample Child module includes questions about the health status of the child and whether a doctor or other health care professional had diagnosed the child with asthma, diabetes, selected developmental conditions, and whether a representative from a school or a health professional had stated that the child had a learning disability.

The health status of sample children aged 0-17 years was asked in the Sample Child HIS section and assessed whether the child's health is generally excellent, very good, good, fair, or poor.

Questions regarding asthma were asked of children aged 0-17 years in the Sample Child AST section and measure the following: whether a doctor or other health care professional ever told that the sample child had asthma; still has asthma; had an episode of asthma or an asthma attack during the past 12 months; and had to visit an emergency room or urgent care center because of asthma during the past 12 months.

Questions regarding diabetes were asked of children aged 0-17 years in the Sample Child DIB section and measure whether a doctor or other health care professional ever told that sample child had prediabetes or borderline diabetes, and diabetes.

Questions regarding development conditions were asked of children aged 0-17 or aged 2-17 years in the Sample Child DLD section. Sample children aged 2-17 years were asked in separate questions whether a doctor or other health professional had ever told that the child had Attention Deficit/Hyperactivity Disorder (ADHD) or Attention-Deficit Disorder (ADD), and autism, Asperger's disorder, pervasive developmental disorder, or autism spectrum disorder. Sample children aged 0-17 years were asked in separate questions whether a doctor or other health professional had ever told that the child had an intellectual disability, also known as mental retardation, and any other developmental delay. For each condition, a follow up question asked whether the sample child currently had this developmental condition.

The DLD section also included two questions for sample children aged 0-17 years that asked whether a representative from a school or a health professional ever told that the child had a learning disability and currently had a learning disability.

### II. Functioning and Disability

#### Annual Core

The questions on functioning and disability found in sections VIS, HEA, MOB, UPP, COM, COG, ANX, DEP, and BEH of the Sample Child module include a Child Functioning Module (CFM) that is part of a set of international standard measures developed, tested and endorsed by the Washington Group on Disability Statistics (WG). The WG is a city group established in 2001 under the United Nations Statistical Commission to address the need for population-based measures of disability by promoting and coordinating international cooperation in the area of health statistics focusing on disability data collection tools suitable for censuses and national surveys. The major

objective is to provide necessary information on disability that is comparable throughout the world by identifying individuals with functional limitations in basic actions, regardless of nationality or culture. The questions reflect advances in the conceptualization of disability and use the World Health Organization's International Classification of Functioning, Disability, and Health (ICF) as a conceptual framework. The CFM was developed jointly with the United Nations Children's Fund (UNICEF). There are questions for children 2-4 years of age (CFM 2-4) and questions for children 5-17 years of age (CFM 5-17). The intended use of these questions is to describe the functional status of children and, when used with other questions on the survey, to evaluate whether children with functional limitations have achieved similar levels of participation and inclusion as children without functional limitations. These questions do not capture all aspects of difficulty in functioning, but rather focus on domains of functioning that are likely to identify the majority of children at risk of participation restrictions in an unaccommodating environment.

Table 9. Functional limitations measured in the Sample Child module annual core, by domain, question topic, questionnaire section and age range of question: 2019 NHIS.

Domain	Question topic	Section	Age range for question
Seeing	Wear glasses or contacts	VIS	2-17
Seeing	Have difficulty seeing (with glasses, if worn)	VIS	2-17
Hearing	Use a hearing aid	HEA	2-17
Hearing	Have difficulty hearing sounds (with hearing aid, if used)	HEA	2-17
Mobility	Use equipment or assistance for walking	MOB	2-17
Mobility	Difficulty walking	MOB	2-4
Mobility	Difficulty walking 100 yards	MOB	5-17
Mobility	Difficulty walking a third of a mile	MOB	5-17
Dexterity	Difficulty picking up small objects	UPP	2-4
Self-care	Difficulty with self-care	UPP	5-17
Communication	Difficulty understanding you	COM	2-4
Communication	Difficulty understanding sample child when they speak	COM	2-4
Communication	Difficulty being understood by people in household	COM	5-17
Communication	Difficulty being understood by people outside of household	COM	5-17
Learning	Difficulty learning things	COG	2-17
Cognition	Difficulty remembering things	COG	5-17
Affect	How often seem very anxious, nervous or worried	ANX	5-17
Affect	How often seem very sad or depressed	DEP	5-17
Playing	Difficulty playing	BEH	2-4
Behavior	Kick, bite, or hit other children or adults	BEH	2-4
Behavior	Difficulty controlling behavior	BEH	5-17
Cognition	Difficulty concentrating	BEH	5-17
Behavior	Difficulty accepting changes in routine	BEH	5-17
Relationships	Difficulty making friends	BEH	5-17

Questions ask about the sample child's level of difficulty (no difficulty, some difficulty, a lot of difficulty, or cannot do at all) in basic domains of functioning including seeing, hearing, mobility, dexterity, self-care, communication, cognition, playing, learning, relationships, and behavior and about the frequency of experiencing anxiety and depression as well as kicking/biting/hitting others. The CFM 2-4 and CFM 5-17 are designed to include domains of functioning relevant for each age group. In addition, several of the domains have questions to provide information on the use of accommodations. Children who use equipment or assistance with walking were asked about difficulty walking both with and without equipment or assistance. For a list of questions asked in each set, and the associated functioning domain, questionnaire section, and respective age range, see Table 9. More information may be obtained by request to the WG Secretariat at [WG\\_Secretariat@cdc.gov](mailto:WG_Secretariat@cdc.gov) or found on the WG website at: <http://www.washingtongroup-disability.com/>.

The WG questions can be analyzed separately, by domain, or combined across domains. CFM disability status indicators are available to data users (see Major Recodes below). The disability composite indicators for sample children aged 2-4 and 5-17 identify children who are at greater risk than the general population for experiencing restrictions in participation because of difficulties doing certain universal, basic actions according to the internationally agreed upon definition developed by UNICEF and the WG. Consult the WG website for guidance on the creation of these indicators. Changing the threshold for either the number of domains the respondent identifies having difficulty with or the degree of difficulty can create different identifiers that will capture different populations. For example, a recode that includes respondents who have "some difficulty" with any of the domains will capture a greater proportion of the population than a recode limited to include only those who report they "cannot do at all" to any of the domains and the functional abilities of the larger group will be much more heterogeneous than that of the smaller group. Analytic guidelines written for each of the CFM questions sets, including recommended disability identifiers, may be obtained from the WG website.

## Major Recodes

**Disability status composite indicator, age 2-4.** An indicator of disability that captures sample children aged 2-4 who are at greater risk than the general population for experiencing restrictions in participation because of difficulties doing certain universal, basic actions according to the internationally agreed upon definition developed by the WG and UNICEF. This recode classifies children with disability as those reporting "a lot of difficulty" or "cannot do at all" for at least one of the questions asking about difficulty seeing, hearing, walking, dexterity, communication, learning, and playing, or who answered "cannot do at all" to controlling behavior. The remaining sample children, that is those reporting "some difficulty" or "no difficulty" to at least one question (and did not report "a lot of difficulty" or "cannot do at all" for any of the questions) are classified as without disability. Those responding "don't know" or "refused" to all questions are excluded.

**Disability status composite indicator, age 5-17.** An indicator of disability that captures sample children aged 5-17 who are at greater risk than the general population for experiencing restrictions in participation because of difficulties doing certain universal, basic actions according to the internationally agreed upon definition developed by the WG and UNICEF. This recode classifies children with disability as those reporting "a lot of difficulty" or "cannot do at all" for at least one of the questions asking about difficulty seeing, hearing, walking, self-care, communication, learning, remembering, concentrating, accepting change, controlling behavior, making friends or who answered "daily" to questions asking how often the sample child feels anxious, nervous, or worried or feels depressed. The remaining sample children, that is those who responded "some difficulty" or "no difficulty" to at least one question (and did not report "a lot of difficulty" or "cannot do at all" for any of the questions) are classified as without disability. Those responding "don't know" or "refused" to all questions are excluded.



### III. Health Care Access and Health Service Utilization

#### Annual Core

Several sections throughout the Sample Child module measure access to and use of health services, as well as affordability of care. Similar content is also available for sample adults. For a list of measures on these topics asked in the annual core, by questionnaire section and reference period of its respective questions, see Table 10.

Table 10. Measures of health care access, service use, and affordability of care asked of the sample child in the annual core, by questionnaire section and reference periods: 2019 NHIS

Measure	Section	Reference period
<b>Immunizations</b>		
Flu vaccine (any, number of vaccines up to 2)	IMM	Last 12 months, Month and Year
<b>Medical Care</b>		
Saw a doctor for medical care	UTZ	Last time interval
Medical and wellness visit combined	UTZ	-
Wellness visit	UTZ	Last time interval
Usual place to go for medical care	UTZ	Current
Kind of place for medical care	UTZ	Current
Number of urgent care visits	UTZ	Last 12 months
Number of emergency department visits	UTZ	Last 12 months
Any overnight hospitalization	UTZ	Last 12 months
Delayed medical care due to cost	UTZ	Last 12 months
Needed but did not get medical care due to cost	UTZ	Last 12 months
<b>Prescription Medication</b>		
Took prescription medication	PMD	Last 12 months
Delayed filling prescription to save money	PMD	Last 12 months
Needed but did not get prescription due to cost	PMD	Last 12 months
<b>Problems Paying Medical Bills</b>		
Anyone in family having problems paying medical bills*	PAY	Last 12 months
Have bills unable to pay at all*	PAY	Current
Level of worry about paying medical bills if sick/accident	PAY	-
*These are family-level replicate questions asked once per family.		

#### Rotating Core

Rotating for two consecutive years every three years starting in 2019 is additional content about sample children's use of selected health services. These questions ask about receiving dental, mental, eye, home, and physical or other therapeutic care, and affordability of dental and mental care. Similar content is also available for sample adults. Table 11 lists rotating core measures of health care use, and affordability, by questionnaire section and reference period of its respective questions.

Table 11. Measures of health care access, service use, and affordability of care asked of the sample child in the rotating core, by questionnaire section and reference periods: 2019 NHIS

Measure	Section	Reference period
<b>Dental Care</b>		
Received a dental cleaning/exam	DNC	Last time interval
Delayed dental care due to cost	DNC	Last 12 months
Needed but did not get dental care due to cost	DNC	Last 12 months
<b>Mental Health Care</b>		
Took medication for emotions/mental health	MHC	Last 12 months
Received therapy or counseling from mental health professional	MHC	Last 12 months
Delayed getting therapy/counseling due to cost	MHC	Last 12 months
Needed but did not get therapy/counseling due to cost	MHC	Last 12 months
<b>Physical and Other Therapeutic/Specialist Care</b>		
Received eye exam from eye specialist	PTC	Last 12 months
Received physical/speech/rehabilitative/occupational therapy	PTC	Last 12 months
Received home care	PTC	Last 12 months

## IV. Behavioral and Mental Health

### Annual Core

The Baby Pediatric Symptom Checklist (BPSC) is a 12-item validated screening tool used for assessing social and emotional difficulties among children aged 0-17 months (Sheldrick, 2013). The BPSC is one component of the larger Survey of Well-being of Young Children (SWYC), a screening instrument designed for use in a clinical setting, such as a pediatric primary care. Information about SWYC is available at <https://www.floatinghospital.org/The-Survey-of-Wellbeing-of-Young-Children/Overview.aspx>

Parents or adults knowledgeable and responsible for the child's health rated a series of behaviors related to irritability, inflexibility, and difficulty with routines that may be used to identify risk for social and emotional difficulties. Each item in this section can be rated as "not at all"; "somewhat"; or "very much" and responses are assigned point values of 0, 1 and 2 respectively. Items with missing responses will count as 0 points. The BPSC is constructed of three subscales (irritability, inflexibility and difficulty with routines) and each subscale is composed of 4 items. Any summed scale of three or more on any of the three subscales indicates that a child is at risk and in practice will prompt further evaluation with a health care professional. For analysis, users may sum responses to operationalize risk as a continuous variable. As thresholds have not yet been developed to operationalize a categorical variable for the general population, no cut-off scores have been provided. Although the BPSC instrument was developed for children under age 18 months, questions were fielded among sample children aged 0-23 months.

### Rotating Core

The Strengths and Difficulties Questionnaire in the SDQ section is a series of questions designed to be used as a screening tool for clinical assessment of mental disorders and for epidemiological research of emotional and behavioral problems in children and adolescents aged 4-17 years. The parent respondent version of the SDQ was



used and this instrument is copyrighted by Robert Goodman, Ph.D., FRCPSYCH, MRCP (Goodman, 1997; Goodman 2009).

Parents or adults knowledgeable and responsible for the child's health were asked to answer, on behalf of the sample child, 25 questions that assess emotional symptoms, conduct problems, hyperactive behavior, peer relationships and prosocial behaviors. The instrument is intended to be scored by summing items within each subscale, and four of the 5 subscales can be summed to create a total score which may be used in analysis as a continuous variable. Higher scores to items assessing emotional symptoms, conduct problems, hyperactive behavior, and peer relationships indicate more problematic attributes. See Table 12 for a list of SDQ subscales by item and scoring values. Recodes of summary scores are available to data users (see Major Recodes below). Additionally, the SDQ section houses questions that collect information about the duration and impact of symptoms, which can be useful for assessing the child's problem, the impact that the problem has on the child and his/her family. Further information about the SDQ, is available at <http://www.sdqinfo.com>.

## Major Recodes

**SDQ subscales.** Summary scores of each of the five SDQ subscales (emotional, conduct, hyperactive, peer relationships and prosocial), have been calculated and provided as a continuous variable ranging from 0-10 with higher scores indicating more problematic attributes for emotional, conduct, hyperactive, and peer relationship subscales and lower scores indicating more problematic attributes for the prosocial subscale. The five subscales can be analyzed separately to look at specific psychological problems. In order to calculate the SDQ subscales, variables with original answer codes of 1, 2, 3 in the instrument which correspond to point values of 0, 1, 2 (when question is worded negatively) or 2, 1, 0 (when question is worded positively) were summed. See Table 12 for subscales with items and point values. SDQ subscale recodes were assigned a code of 88 if more than two items needed to calculate the subscale were answered as "don't know," "refused," or were not ascertained.

**SDQ total.** The SDQ total score is based on the sum of the first four subscales (emotional, conduct, hyperactive, and peer relationships). This recode is a continuous variable ranging from 0-40 with higher scores indicating more problematic attributes. For the SDQ total score, a code of 88 has been assigned if more than one subscale score is missing.

**SDQ impact score.** The SDQ impact score is based on the sum of the impact questions, which are asked only among sample children 4-17 years who had responded as having had minor, definite or severe difficulties with emotions, concentration, behavior, or being able to get along with other people. This recode is a continuous variable ranging from 0-10 with higher scores indicating that difficulties have a greater impact on the child's and or family's life. For the SDQ impact total score a code of 88 has been assigned if more than one item is missing.

**Prosocial subscale.** A fifth subscale that describes children's positive behaviors. These items are excluded from the overall SDQ scoring when using the SDQ to identify children who have or are at increased risk of having psychological disorders.

Table 12. Strengths and Difficulties Questionnaire subscales, items in each subscale, and point values for scoring

Subscale and Question Topic	Not true	Somewhat true	Definitely true
<b>EMOTIONAL SUBSCALE</b>			
Often complains of headaches, stomach-aches, or sickness	0	1	2
Many worries, often seems worried	0	1	2
Often unhappy, depressed or tearful	0	1	2
Nervous or clingy in new situations	0	1	2
Many fears, easily scared	0	1	2
<b>CONDUCT SUBSCALE</b>			
Often has temper tantrums or a hot temper	0	1	2
Generally obedient, usually does what parents want	2	1	0
Often fights with other children or bullies	0	1	2
Often lies or cheats	0	1	2
Steals from home, school, or elsewhere	0	1	2
<b>HYPERACTIVITY SUBSCALE</b>			
Restless, overactive, cannot stay still for long	0	1	2
Constantly fidgeting or squirming	0	1	2
Easily distracted, concentration wanders	0	1	2
Thinks things out before acting	2	1	0
Sees task through to the end, good attention span	2	1	0
<b>PEER RELATIONSHIP SUBSCALE</b>			
Rather solitary, tends to play alone	0	1	2
Has at least one good friend	2	1	0
Generally liked by other children	2	1	0
Picked on or bullied by other children	0	1	2
Gets on better with adults than other children	0	1	2
<b>PROSOCIAL SUBSCALE</b>			
Considerate of other people's feelings	0	1	2
Shares readily with other children	0	1	2
Helpful if someone is hurt, upset, or feeling ill	0	1	2
Kind to younger children	0	1	2
Often volunteers to help others	0	1	2

Source: Goodman R. The Strengths and Difficulties Questionnaire: a research note. *Journal of child psychology and psychiatry, and allied disciplines*. 38(5):581-6. 1997.

---

## V. Stressful Life Events

### Rotating Core

Stressful life events have been associated with early and lifelong health conditions that may extend into adulthood, including diabetes, mental health disorders, and cardiovascular problems (Felitti et al., 1998). The NHIS includes four questions in the SLE section pertaining to stressful life events, including whether the child has witnessed or experienced neighborhood violence, lived with someone with a mental illness, lived with someone with a drug or alcohol problem, or had a parent who was incarcerated after the child was born. These questions were adapted from the set of eight questions on stressful life events appearing in the National Survey of Children's Health (NSCH). The NHIS does not include the full set of questions, as other stressful life events can be ascertained through other questions asked within the NHIS or were deemed inappropriate to ask in a face-to-face interview. As result, there is not a standard way of analyzing the questions in this section, nor do they need to be combined to form a composite.

Given the sensitive nature of these questions, an introductory paragraph precedes the first question and informs the respondent that these events can happen in any family and that any question can be skipped. It is possible that some respondents may choose to skip some of these questions, particularly if other individuals in the household are present at the time of the interview.

---

## Health Insurance

### Annual Core

The health insurance sections (INS) of the Sample Adult and Sample Child modules have a full range of items addressing health insurance such as coverage status, sources of coverage, characteristics of coverage, and reasons for no coverage. The flow and content of the questions pertaining to health insurance programs covered in the INS sections are similar to questions covered in the 1997-2018 NHIS Family Core. The main difference is that instead of asking health insurance for all family or household members, one adult and one child (if present) are selected from each household to receive these questions. The sample adult and sample child receive a similar set of questions with a few exceptions that will be outlined below.

### Health Insurance Coverage Status

An individual is considered currently insured if they currently have coverage through private health insurance, Medicare, Medicaid, Children's Health Insurance Program (CHIP), military (TRICARE, Veterans Administration (VA), and CHAMP-VA), other state-sponsored health plans, or other government program. Individuals without any of the aforementioned coverages or with only Indian Health Service coverage or a non-comprehensive plan that covers only dental, vision, or prescription drugs are considered uninsured.

For ease of analysis two recodes are available, NOTCOV\_A (on the Sample Adult file) and NOTCOV\_C (on the Sample Child file) that reflect this definition of noncoverage as used in *Health, United States* (in which persons with *only* Indian Health Service coverage or a single service plan that covers only dental, vision, or prescription drugs are considered uninsured).

### Sources of Coverage

Sample adult and sample child respondents could identify one or more sources of medical care coverage, and single service plans were asked as separate questions. The following sources of healthcare coverage were collected in the interview:

- Private health insurance: Coverage obtained through employment or directly purchased (including Medigap plans)
- Medicare: The federal health insurance program for adults who are 65 and older, certain younger people with disabilities, and people with End-Stage Renal Disease (permanent kidney failure requiring dialysis or a transplant, sometimes called ESRD)
- Medicaid: A joint federal and state program that provide free or low-cost healthcare coverage to Americans, including some low-income people, families and children, pregnant women, the elderly, and people with disabilities
- CHIP (Children's Health Insurance Program): A joint federal and state program that provides low-cost health coverage to children in families that earn above the income threshold to qualify for Medicaid

- Military: TRICARE, VA or CHAMP-VA
- Other state-sponsored health plans
- Other government program
- Indian Health Service: A part of the federal government that delivers direct medical and public health services to federally recognized Native American Tribes and Alaska Native people
- Single service dental plan
- Single service vision plan
- Single service prescription drug plan

### *Characteristics of Coverage*

For all coverage types, except for the Indian Health Service, additional follow-up questions specific to the type of coverage are asked. Some of these characteristics of coverage are broad and are relevant to more than one coverage type, whereas other characteristics are very specific to a particular type of coverage. In summary, the INS section includes detailed follow-up questions in the following areas:

- Health Insurance Marketplace, state exchanges or Healthcare.gov as to how coverage (private, Medicaid, CHIP, state-sponsored plans, other government programs) was obtained
- Enrollment in a high deductible health plan (private, Medicaid, CHIP, state-sponsored plans, or other government programs)
- Premium (private, Medicaid, CHIP, state-sponsored plans, or other government programs)
- Characteristics of private insurance (up to two plans per person)
  - exchange-based
  - policyholder
  - coverage of individuals other than the policyholder
  - relationship to the policyholder (only sample adult)
  - how plan was obtained (e.g. work, directly purchased)
  - who pays for the plan (e.g. self or family, work)
  - annual amount of premium paid by individual or family
  - prescription drug benefit
  - dental benefit
  - vision benefit
  - health savings account
- Characteristics of Medicare
  - Medicare Parts (A, B, C, D)
  - Medicare Advantage plan or Medicare managed care

- Types of military healthcare
  - VA (only sample adult)
  - TRICARE
  - CHAMP-VA

### *Continuity of Coverage*

For persons with coverage, they were asked if there was any time in the past 12 months when they did not have coverage. If they answered “yes,” they were asked for how many months they did not have coverage. For persons who were currently uninsured, they were asked when the last time was that they had coverage. If it was less than a year, they were asked for how many months they did not have coverage.

### *Reasons for No Coverage*

There are two sets of questions concerning the reasons for not having healthcare coverage. The first set focuses on reasons for no longer being enrolled in health coverage among those individuals who currently lacked coverage for less than three years. Reasons include: policyholder retired, lost a job or changed employers, a deadline was missed for signing up or paying for coverage, became ineligible due to age or leaving school, the cost of the coverage increased, and had Medicaid or other public coverage but were no longer eligible. The second set of questions focuses on reasons for not obtaining health coverage among all individual who currently did not have healthcare coverage. Reasons include: currently uninsured because coverage is not affordable, do not need or want coverage, not eligible for coverage, process of signing up is too difficult or confusing, cannot find a plan that meets needs, has applied for coverage but it has not started yet and other reason. Based on coding open-ended responses, two additional categories are available as recoded variables. These additional reasons captured were retired, lost a job or changed employers and missing a deadline for signing up for coverage.

### *Replicate*

To reduce respondent burden, under certain circumstances households who shared the same private plans were only asked about detailed characteristics of shared plans once, either in the Sample Adult interview or Sample Child interview, whichever occurred first. To be eligible for this replicate, the sample adult and sample child must be from the same family, the private plan must cover more than one person, the private plan must have a “valid” plan name, (i.e. it cannot have a refused or not known as the name of the plan), and the plan has to have information as to either where the plan was obtained or who pays for it, (i.e. it cannot have refused or not known to either of these fields). In addition, if either the sample adult or sample child indicated that they have Medigap coverage through HIKIND03\_A or HIKIND03\_C, they were not eligible for the replicate. Families may share up to two private plans.

### *Processing Health Insurance Responses*

The INS sections use responses to follow-up questions to evaluate the reliability of the reported health insurance coverage and to adjudicate conflicting information. For many survey respondents, health insurance is a complex topic and some inconsistencies in survey responses are expected. If the responses to follow-up questions are inconsistent with the original health insurance coverage indicated, the original responses are edited. As a result, a portion of the sample adults and sample children are reassigned to a different type of coverage or reclassified from insured to uninsured (or vice versa). Conversely, follow-up responses in agreement with the original health insurance response are not edited and are included in the recodes. Therefore, it is best to use the recodes created, and listed in Table 13 below, for specific types of healthcare coverage and noncoverage because of the complicated editing process that takes place in the INS sections.

Table 13. List of health insurance recode variables for sample adult and sample children: 2019 NHIS

<b>Type of health insurance coverage</b>	<b>Sample Adult file</b>	<b>Sample Child file</b>
Private health plans	PRIVATE_A	PRIVATE_C
Medicare	MEDICARE_A	OTHGOVR_C*
Medicaid	MEDICAID_A	MEDICAID_C
Children's Health Insurance Program (CHIP)	CHIP_A	CHIP_C
Military health plans	MILITARY_A	MILITARY_C
Indian Health Service	IHS_A	IHS_C
Other government programs	OTHGOV_A	OTHGOVR_C*
State-sponsored health plans	OTHPUB_A	OTHPUB_C
Uninsured	NOTCOV_A	NOTCOV_C
*This recode combines sample children covered by Medicare, other government programs or both.		

## Characteristics about the Sample Adult and Sample Child

### Annual Core

#### *Sex, Age, Hispanic origin and Race*

The 2019 NHIS collected information across different modules regarding variables that describe the sample adult and sample child's sex, age, Hispanic origin, and race. Table 14 lists the variables that summarized the final public use variables on sex, age, Hispanic origin and race available for the sample adult and sample child. These variables are described in the HHC section of the Codebooks.

Table 14. List of public use sex, age, Hispanic origin and race variables: 2019 NHIS

Description	Sample adult variable	Sample child variable
Sex	SEX_A	SEX_C
Age	AGEP_A	AGEP_C
Hispanic origin	HISP_A	HISP_C
Hispanic group detail	HISDETP_A	HISPDETP_C
Single and multiple race groups	RACEALLP_A	RACEALLP_C
Single and multiple race groups combined with Hispanic origin	HISPALLP_A	HISPALLP_C

Starting in 2019, responses of “refused” or “don’t know” to the sex and age questions are allowed. To preserve confidentiality, the ages of adults aged over 85 years are top-coded at 85 on the Sample Adult public use data file.

In accordance with the Office of Management and Budget’s Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity (OMB, 1997) for the collection of ethnicity and race in federal data systems, separate questions are asked about Hispanic origin and race. Persons of Hispanic origin may be of any race or combination of races. Hispanic origin includes persons of Mexican/Mexican American/Chicano, Central American, South American, Puerto Rican, Cuban, Dominican, or other Hispanic origin. Race is based on the sample adult’s description of his or her own racial and ethnic identity, and an adult knowledgeable and responsible for the child’s health provides this information for the sample child. More than one race can be reported for a sample adult and sample child.

The confidentiality of respondents and their families could be compromised if there were extensive details available about the sample adult, sample adult’s spouse or partner, sample children and their parents, and other family members. Starting in 2019, detailed information about race and ethnicity, that was previously available in public use files, is suppressed as more detailed geographic information (e.g. region and urbanization level) is included in the public use file. The following four single race categories are available for sample adult and sample children in the public use files: 1) white; 2) black or African American; 3) Asian; and 4) American Indian or Alaska Native (AIAN). The only multiple race category available in the public use files is AIAN and another race.



Sample adult and sample child respondents indicating a single race other than four mentioned or reporting more than one race, other than including AIAN, were combined into “other single and multiple races” category. Detailed Hispanic origin information available in the public use files includes Mexican and Mexican American only.

For additional information about the historical context of race and ethnicity data collection in the NHIS, including editing, references for NHIS race and Hispanic origin, and background documents related to race and ethnicity data collection in federal data systems, see Appendix II of the 2012 NHIS Survey Description.

### *Nativity*

Information on place of birth (U.S. state or territory, or outside of the U.S.) and citizenship status is collected in the NAT section for Sample Adults and Sample Child modules. To protect confidentiality, information on U.S. state or territory of birth, and detailed citizenship status are not available in the public use data files. Information about whether the sample adult or sample child was born in the U.S., and for those born outside the U.S., whether the sample adult or sample child is a U.S. citizen and years in the U.S. (categorical variable) is available in the public use files.

### *Schooling*

The schooling sections (SCH) of the Sample Adult and Sample Child modules contain questions to determine the number of school days missed during the past 12 months. All sample children under aged 5–17 years, and sample adults aged 18 and over who were enrolled in or attending school at the time of interview were asked about the number of school days missed because of illness, injury, or disability during the 12 months prior to the interview. Persons responding for the sample children aged 0–17 years were asked if the child ever had a special education or early intervention plan, such as an Individualized Education Plan (IEP), or an Individualized Family Service Plan (IFSP). Those who responded “yes” were then asked if the child currently (sometime in the past school year) a special education or early intervention plan, and if he/she received these services to help with his/her emotions, concentration, behavior, or mental health.

### *Education*

Information on educational attainment, measured as the highest level of school or highest degree completed is asked about all adults in the household during the household roster module. The variable EDUC\_A describes the highest educational attainment of the sample adult. Additionally, recodes have been created that describe the level of educational attainment for the spouse or partner of the sample adult (described under “Characteristics about the Spouse or Partner of the Sample Adult”), and the parents of the sample child (described under “Characteristics about the Sample Child’s Parents”). Sample children are not asked about highest school grade or education completed.

### *Employment*

The Sample Adult EMP section contains information regarding the sample adult’s work status in the week before the interview, main reason for not working, hours worked, work benefits for those working, and days of sick leave taken among sample adults who worked in the past year.

Sample adults were first asked whether they worked for pay at a job or business last week; if not, they were asked if they had a job or business last week, but were temporarily absent due to illness, vacation, family or maternity leave. Those who were working or temporarily absent from work were then asked how many hours in total they worked at all jobs or business in the last week (if they said they were working last week) or in a usual week (if they were temporarily absent last week). Sample adults who said they were working last week and said they worked 34 or fewer hours are asked if they *usually* work 35 or more hours per week in total at all their jobs or businesses.

Sample adults who were not working last week or were not temporarily absent from a job or business last week are asked the main reason they were not working for pay at a job or business last week. Sample adults who couldn't find work, were laid off, looking for work, retired, unable to work for health reasons/disabled, taking care of house or family, going to school, or some other reason, as well as refused or don't know responses are then asked when was the last time they worked for pay at a job or business, even if only for a few days.

Employed sample adults – those who were working last week, temporarily absent last week, performed seasonal or contract work, or were working, but not for pay – are asked whether paid sick leave is available and whether health insurance is offered through their workplace. Lastly, all sample adults working within the past 12 months are asked how many days of work they missed because of illness, injury or disability in the past year. To protect confidentiality, information on the number workdays missed due to health reasons is top-coded at 130 days in the recode EMPDYSMSS2.

### ***Employment Recodes***

Sample adults who performed seasonal or contract work or worked, but not for pay, in a job or business were considered employed and were asked about paid sick leave, health insurance available through the workplace, and work days missed in the last 12 months even though they indicated that they were not working in the past week and were not temporarily absent from work. In addition, they were not asked about the number of hours they worked in the last week or in a usual week and were not asked the last time they worked for pay. To insure that the variables in this section are internally consistent with one another, a summary recode, EMPWRKLSWK\_A, identifies all employed sample adults, including those who were working during the last week; those with a job or business but temporarily absent the last week; and those performing seasonal, contract, or unpaid work. A second recode, EMPWKHRS2\_A, provides the hours worked last week (top-coded at 95 for confidentiality), and codes sample adults who performed seasonal, contract, or unpaid work as “998” or “not ascertained” on this recode. A third recode, EMPWRKFT\_A, identifies all employed sample adults who usually worked 35 or more hours per week; those performing seasonal, contract, or unpaid work are coded as “8” or “not ascertained” on this recode. Lastly, a fourth recode, EMPLSTWORK\_A, indicates the last time that sample adults who were not working in the last week or temporarily absent from their usual job or business had worked for pay, and sample adults performing seasonal, contract, or unpaid work are coded as “8” or “not ascertained” on this recode.

### ***Marital Status***

A series of questions in the MAR section collects information regarding the marital status of sample adults. Sample adults are first asked if they are “now married, living with a partner together as an unmarried couple, or neither.” Married sample adults are asked if their spouse lives in the same residence; if not, they are asked if this is because the sample adult and his or her spouse are legally separated. Sample adults are also asked to verify the sex of their spouse or partner that was obtained during rostering, and to correct it, if necessary. Sample adults who are living with an unmarried partner or who are neither married nor living with a partner or don't

know or refuse to state their marital status are asked if they have ever been married. Sample adults who are currently living with a partner and have been married are asked their current legal marital status – that is, whether they are currently married, widowed, divorced, or separated. Sample adults who are neither living with a partner nor married but have been married are asked if they are widowed, divorced, or separated. Additional information about the spouse or partner of the sample adult, also collected in the MAR section, is available under “Characteristics about the Spouse or Partner of the Sample Adult.”

### *Sexual Orientation*

Sample adults were asked about their sexual orientation. This question was asked before determining marital status of the sample adult and sex of the spouse or partner living in the household. Sexual orientation was not asked about the sample child.

### *Proxy Status*

Generally, sample adults provide information for themselves during the Sample Adult interview. However, in a small number of cases, proxy responses are allowed if the sample adult had a physical or mental condition that prevented them from responding. The variable PROXY\_A indicates those cases for which a proxy respondent provided the information.

### *Veteran Status*

The VET section contains information about the sample adult’s military veteran status and use of VA services. Specifically, it includes information on whether the sample adult ever served in the U.S. Armed Forces, military Reserves, or National Guard, if served in active duty for training or in a combat setting or humanitarian peace-keeping mission, and whether has received a service-connected disability rating. Information about use of VA services focused on whether the sample adult sought care from at VA Hospital or other VA-affiliated facility in the past 12 months, and whether they have ever enrolled in or used VA healthcare.

---

## Characteristics about the Parents of the Sample Child

### Annual Core

Starting in 2019, the NHIS includes several demographic characteristics about parents residing with the sample child. The Child PAR section asks an adult knowledgeable and responsible for the child's health responding on behalf of the sample child a series of questions about the type of relationship (biological, adoptive, step, foster) between the child and his or her parent(s), and each parent's marital status and birthplace (whether the parent was born in the U.S. or a U.S. territory). The redesigned NHIS allows for the collection of information of up to four parents as long as they all reside in the same household as the child.

Some detailed relationship information is suppressed in the public use data due to disclosure risks. Specifically, adoptive and biological children are combined in a single category, and foster children cannot be identified. In addition, only a very small number of sample children in the 2019 data were residing with three parents, and there were no instances where a sample child had four residential parents.

Several recodes are available on the public use data that describe the demographic characteristics of up to two parents residing with the sample child. The information for these recodes was obtained from questions asked in various sections of the Sample Child module (i.e., PAR, GEN, HHC), and are available in the PAR section of the Sample Child module. Table 15 lists the parental recodes available to data users which include parental sex, age (bottom-coded at 20 and top-coded at 65), education of the parent with the highest educational attainment, employment status (including full- versus part-time work), marital status and legal marital status, and whether the sample child parents are of the same sex or of opposite sex. Note that parents currently serving in the U.S. Armed Forces cannot be included in the public use data because of disclosure risks.

To protect confidentiality, detailed information about race and ethnicity of the sample child's resident parents is suppressed. To assist data users interested in studying interracial families, the PAR section also includes four recodes to indicate (1) whether the sample child's race is the same as all parents in the household, (2) whether the sample child and all the parents in the household are of same Hispanic or Latino origin category (i.e. Hispanic, non-Hispanic), (3) whether the sample child's parents are of the same race to each other, and (4) whether the sample child's parents are of the same Hispanic or Latino origin category to each other. These recodes are "yes" and "no" answers. The "yes" response means that the sample child and all resident parents are in the same racial category, or the sample child and all resident parents are in the same Hispanic or Latino origin category. A "no" response means that either the sample child and at least one of the parents is of a different racial category, or that either the sample child and at least one of the parents are not of the same Hispanic or Latino origin category, respectively. Similarly, same race and same Hispanic or Latino origin category between parents is based on a common racial and Hispanic or Latino origin category. Same race is in reference to the racial categories available in the public use file, that is white, black or African American, Asian, AIAN, AIAN and another race, and all other single and multiple races. Same Hispanic or Latino origin category is based on whether the child and all parents, and whether all the parents are of any Hispanic or Latino ethnicity (regardless of country or area of origin). Sample children residing with one parent in the household are categorized as missing (not in universe) in the variables categorizing whether the sample child's parents are of the same race to each other, and whether the sample child's parents are in the same Hispanic or Latino origin category.

Additional information about the family of the sample child is described under "Characteristics about the Family and Household of the Sample Adult and Sample Child."

Table 15. Public use recodes describing demographic characteristics of the parents residing with the sample child: 2019 NHIS

Description	Codebook section	Residential parent 1 variable	Residential parent 2 variable	Variable for both residential parents
Sex	PAR	PARSEX1_C	PARSEX2_C	
Age	PAR	PARAGETC1_C	PARAGETC2_C	
Education of the sample child's parent with the highest education	PAR			MAXPAREduc_C
Working last week	PAR	PARWORK1_C	PARWORK2_C	
Working full-time last week (35+ hours)	PAR	PARWKFT1_C	PARWKFT2_C	
Current marital status	PAR	MARSTAT1_C	MARSTAT2_C	
Legal marital status	PAR	LEGMSTAT1_C	LEGMSTAT2_C	
Two parents of same or of opposite sex	PAR			PARSAMEOPP_C
Sample child's Hispanic ethnicity is the same as both parents	PAR			SCPARHISP_C
Sample child's race is the same as both parents	PAR			SCPARRAC_C
Sample child's parents are of the same Hispanic ethnicity	PAR			HISPPARSC_C
Sample child's parents are of the same race	PAR			RACPARSC_C

Note: The sample child's parents are labeled as parent 1 and parent 2 based in the order that this information was provided by the respondent and entered by the interviewer during the interview for questions WHOPAR and WHOFOST. All information in the recodes about parent 1 (e.g. sex, age) are about the same person, and similarly for parent 2.

## Characteristics about the Spouse or Partner of the Sample Adult

### Annual Core

Starting in 2019, the NHIS includes several recodes indicating demographic characteristics of the spouse or partner living with the sample adult, if married or cohabiting. The information for these recodes was obtained from questions asked in various sections of the survey (i.e., MAR, GEN, HHC), and are available in the MAR section of the Sample Adult module. These recodes include the spouse or partner's sex, age (bottom-coded at 20 and top-coded at 85 years), education, current work status, and whether the spouse or partner worked full-time (35 or more hours per week) and are shown in Table 16.

For data users interested in studying interracial families, recodes of the race and ethnicity of the spouse or partner in the household are available in reference to the sample adult's race and ethnicity. These recodes are "yes" and "no" answers, where a "yes" response means that sample adult and the spouse or partner are in the same racial category, or the same Hispanic or Latino origin category (i.e. Hispanic, non-Hispanic), and a "no" response means that the sample adult and the spouse or partner are of a different racial category, or Hispanic or Latino origin category, respectively. Same race is in reference to the racial categories available in the public use file, that is white, black or African American, Asian, AIAN, AIAN and another race, and all other single and multiple races. Same Hispanic ethnicity is based on any Hispanic or Latino ethnicity (regardless of country or area of origin).

Additional information about the family of the sample adult is described under "Characteristics about the Family of the Sample Adult and Sample Child."

Table 16. Public use recodes describing demographic characteristics of the spouse or partner residing with the sample adult: 2019 NHIS

Description	Codebook section	Variable for spouse	Variable for cohabiting partner
Sex	MAR	SPOUSESEX_A	PRTNRSEX_A
Age	MAR	SPOUSAGETC_A	PRTNRAGETC_A
Hispanic ethnicity of sample adult and spouse/partner are the same	MAR	SASPPHISP_A	SASPPHISP_A
Race of sample adult and spouse/partner are the same	MAR	SASPPRACE_A	SASPPRACE_A
Education	MAR	SPOUSEDUC_A	PRTNREDUC_A
Working last week	MAR	SPOUSWRK_A	PRTNRWRK_A
Working full-time last week (35+ hours)	MAR	SPOUSWKFT_A	PRTNRWKFT_A

## Characteristics about the Family and Household of the Sample Adult and Sample Child

### Annual Core

#### Family and Household Composition variable

Table 17 provides a list of various recodes included on the public use Sample Adult and Sample Child data files that describe the families and households in which the sample adult and sample child live. Flag variables indicate source respondent for each module and whether the sample unit was single or multiple family household. Also listed are recodes that indicate the level of education attained by the adult with the highest education in the sample adult's and sample child's family. In addition, several top-coded counters provide the number of household members in the sample adult's and sample child's family, and family members in various age groups.

Table 17. Public use recodes of family and household composition and counter variables available for sample adult and sample child: 2019 NHIS

Description	Codebook section	Variable name for sample adult's family	Variable name for sample child's family
Sample adult is the household respondent or the proxy who lives in the household	FLG	HHRESPSA_FLG	
Sample child respondent is the household respondent	FLG		HHRESPSC_FLG
Number of persons in the sample adult's /sample child's family (top-coded)	FAM	PCNTFAM_A	PCNTFAM_C
Number of adults in the sample adult's / sample child's family (top-coded)	FAM	PCNTADLT_A	PCNTADLT_C
Number of children in the sample adult's / sample child's family (top-coded)	FAM	PCNTKIDS_A	PCNTKIDS_C
Indicator for at least one person is 65 and over in the sample adult's / sample child's family	FAM	OVER65FLG_A	OVER65FLG_A
Education of the adult with the highest education in the sample adult's family / sample child's family	FAM	MAXEDUC_A	MAXEDUC_C
Flag indicating sample adults / sample child living in households containing more than one family	FAM	MLTFAMFLG_A	MLTFAMFLG_C
Counters of all members in the sample adult's and sample child's households (top-coded)	HHC	PCNTTC	PCNTTC
Counters of persons aged 0-17 years in the sample adult's and sample child's households (top-coded)	HHC	PCNTLT18TC	PCNTLT18TC
Counters of persons aged 18 years and older in the sample adult's and sample child's households (top-coded)	HHC	PCNT18UPTC	PCNT18UPTC

For 140 households in 2019, there were inconsistencies in the data that suggested the household roster was incomplete. The number of household members in the sample adult and sample child's family is unknown. Counter variables (that is, the person count variables for the household and Sample Adult and Sample Child families) have been set to 8 (not ascertained).

### *Family Employment*

The 2019 NHIS family employment modules (FEM) obtained employment information for all related adults in the sample adult's and sample child's families. To reduce respondent burden, these questions were asked once per family. Responses to these questions were used to create several counters on the public use sample adult and sample child file and are top-coded for confidentiality, see Table 18.

Table 18. Employment counter variables available in the sample adult or sample child files: 2019 NHIS

Description	Module, section	Variable Name	Notes
Number of adults in the sample adult's family who are working	Sample Adult, FEM	PCNTADTWKP_A	Top-coded at 3.
Number of adults in the sample adult's family who are working full-time (35 or more hours per week)	Sample Adult, FEM	PCNTADTWFP_A	Top-coded at 3.
Number of adults in the sample child's family who are working	Sample Child, FEM	PCNTADTWKP_C	Top-coded at 3.
Number of adults in the sample child's family who are working full-time (35 or more hours per week)	Sample Child, FEM	PCNTADTWFP_C	Top-coded at 3.

### *Family Income*

The family income section (INC) contain information regarding a variety of income sources, as well as estimates of total combined family income. All questions are asked once per family, using the family-level-replicate interviewing approach. Respondents are told at the start of the family income section that all questions are seeking information about possible income sources in the previous calendar year, and the names of all family members (collected earlier) to consider when responding.

Respondents were asked whether anyone in the family received income from a variety of sources (e.g. wages, salary from self-employment, social security, railroad retirement, government assistance). Respondents are also asked to report their "best estimate" of their family's total income (in dollars) from all sources for all family members living in the household before taxes in the last calendar year. Because nonresponse to this question tends to be relatively high, the NHIS includes a series of follow-up questions utilizing an unfolding bracket methodology that obtains additional income information. The unfolding bracket method asked a series of closed-ended income range questions (e.g., "is it less than \$75,000, or \$75,000 or more?") if the respondent did not provide an estimated total family income. These closed-ended income range questions were constructed so that each successive question established a smaller range for the amount of the family's income. In addition to



asking respondents about the family's income relative to specific dollar values (i.e., \$75,000, \$100,000, and \$150,000), these respondents were also asked about the family's income relative to the federal poverty threshold (100%, 138%, 200%, 250% and 400%) and take into account each family's size (collected earlier in the interview).

The poverty thresholds used in the questionnaire, and shown in Table 19, were estimated from several sources: the weighted average poverty thresholds for the previous year from the U.S. Census Bureau; and from the U.S. Bureau of Labor Statistics, the average Consumer Price Index from two years ago (actual Consumer Price Index values for January–July, and projected Consumer Price Index values for August–December).

Table 19. Poverty thresholds used in the 2019 NHIS Instrument, by family size.

Family Size	100% of the federal poverty level	138% of the federal poverty level	200% of the federal poverty level	250% of the federal poverty level	400% of the federal poverty level
1 person < 66 years	\$13,000	\$18,000	\$26,000	\$33,000	\$52,000
1 person ≥ 66 years	\$12,000	\$17,000	\$24,000	\$30,000	\$48,000
2 persons, both < 66	\$17,000	\$23,000	\$34,000	\$42,000	\$68,000
2 persons, 1 is ≥ 66	\$15,000	\$21,000	\$30,000	\$38,000	\$61,000
3 persons	\$20,000	\$28,000	\$40,000	\$50,000	\$80,000
4 persons	\$26,000	\$35,000	\$51,000	\$64,000	\$103,000
5 persons	\$30,000	\$42,000	\$61,000	\$76,000	\$122,000
6 persons	\$34,000	\$48,000	\$69,000	\$86,000	\$138,000
7 persons	\$39,000	\$54,000	\$78,000	\$98,000	\$157,000
8 persons	\$44,000	\$60,000	\$87,000	\$109,000	\$175,000
9 or more persons	\$52,000	\$72,000	\$104,000	\$130,000	\$208,000

When the questions about income relative to poverty threshold are asked during the course of the interview, the appropriate poverty threshold relative to the family's size (in a dollar amount) is displayed on the interviewer's screen, so that the respondent is asked if the family's income in the previous year was less than the applicable NHIS poverty threshold, or if the family's income was greater than or equal to that same poverty threshold.

To protect confidentiality, family income reported in dollar amounts as well as the variables obtained from the income bracketing questions are not available on the NHIS public use data files. See Appendix for availability of restricted income questions.

### *Recodes of Family Income and Imputed Family Income*

Missing data on family income and earnings in the NHIS are imputed using a multiple imputation methodology. Imputation is the process of replacing missing data with substituted values based on information collected from

other observations in the dataset. Multiple imputation accounts for the extra variability due to imputation in statistical analyses.

Starting in 2019, two separate files are available as part of the data release: `adultinc19` and `childinc19` which contain 10 multiply imputed income data values for the sample adult and the sample child's families, respectively. All 10 imputations are stacked in a single file with a variable (`IMPNUM`) indicating the imputation number. In previous years, NHIS core data files were accompanied by five imputations of selected income data in five separate datasets. However, recent literature on multiple imputation analysis suggests that increasing the number of imputations (e.g. to 10 or higher) produces more precise estimates for a wide variety of analyses (van Buuren, 2012). Stacking the 10 multiple imputation datasets into one allow for fewer steps in data preparation for analyses in SAS and Stata. See the examples provided further in this section.

Ten sets of top-coded variables for family income, poverty ratio, grouped income and grouped poverty ratios are available for the sample adult and sample child families in the two income files respectively (Table 20). These recodes incorporate information from reported and imputed total family income and are available as a continuous total family income value and in categories of family income, and as the continuous ratio of total family income and family size relative to the poverty threshold and as a finite number of categories.

Table 20. Top-coded variables for family income and poverty ratio, and related flags available in the Sample Adult and Sample Child files

Description	Sample Adult file	Sample Child file
Top-coded family income	FAMINCTC_A	FAMINCTC_C
Grouped family income	INCGRP1_A	INCGRP_C
Top-coded poverty ratio	POVRATTC_A	POVRATTC_C
Grouped poverty ratio	RATCAT_A	RATCAT_C
Income top-code flag	INCTCFLG_A	INCTCFLG_C
Imputed income flag	IMPINCFLG_A	IMPINCFLG_C

In cases where the sample adult and sample child are in the same family, these corresponding values are identical. The family income is top-coded at the 95<sup>th</sup> percentile and imputed within the lower and upper bound when the income bracketing questions are answered. Similarly, the poverty ratio variable is top-coded at the 95<sup>th</sup> percentile.

For the convenience of analyses that don't need or use multiple imputed data, the same variables, i.e., `FAMINCTC_A`, `INCGRP_A`, `POVRATTC_A`, `RATCAT_A`, and `INCTCFLG_A` (for the sample adult's family) and analogously for the sample child family from a single imputation are also available in the Sample Adult and Sample Child files, respectively. Descriptive statistics for the continuous variables for each imputation are available in the appendix of the codebook.

No personal earnings information is collected as part of the redesigned questionnaire.

For technical information about the imputation model, please refer to the "Imputed Income Technical Document" available with the 2019 file releases on the NHIS website, under "Using the NHIS." For examples of code on how to use multiply imputation in analysis, see the "Merging files" section in this document.

---

## *Food Related Programs*

The food related programs (FOO) section includes three questions to ascertain past 12-month participation in the food assistance program Supplemental Nutrition Assistance Program (SNAP), free or reduced-cost breakfasts or lunches at school, and the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC). These are family-level replicate questions asked once per family. The universes for the questions in this section coincide with program eligibility. Sample adult and sample child respondents were asked about SNAP assistance program participation. Sample adult and sample child respondents living in families with females 12-55 years of age or children 0-5 years of age were asked about WIC program participation. Sample adult and sample child respondents living in families with children between the ages of 5-17 were asked about free or reduced-cost breakfasts or lunches at school.

## *Housing*

The housing section (HOU) collects information on housing tenure, length of residence, and participation in Federal, State, or local government housing assistance programs among renters. To reduce respondent burden, these questions were asked once per family.

## *Region and urbanization level*

Geographical classification of the U.S. population is provided on the NHIS in two ways: region and urban-rural classification. In the geographical classification, states are grouped into four regions used by the U.S. Census Bureau: northeast (Maine, Vermont, New Hampshire, Massachusetts, Connecticut, Rhode Island, New York, New Jersey, and Pennsylvania), Midwest (Ohio, Illinois, Indiana, Michigan, Wisconsin, Minnesota, Iowa, Missouri, North Dakota, South Dakota, Kansas, and Nebraska), south (Delaware, Maryland, District of Columbia, West Virginia, Virginia, Kentucky, Tennessee, North Carolina, South Carolina, Georgia, Florida, Alabama, Mississippi, Louisiana, Oklahoma, Arkansas, and Texas), and west (Washington, Oregon, California, Nevada, New Mexico, Arizona, Idaho, Utah, Colorado, Montana, Wyoming, Alaska, and Hawaii).

The urban-rural classification is based the 2013 NCHS Urban-Rural Classification Scheme for Counties, which groups U.S. counties and county-equivalent entities into six urban-rural categories: large central metro, large fringe metro, medium metro, small metro, micropolitan, and non-core (Ingram and Franco, 2014). The categorization included on the NHIS public use files combines medium and small metropolitan areas into a single group and micropolitan and non-core areas into a single group to yield the following four categories: large central metro, large fringe metro, medium and small metro, and nonmetropolitan. Additional information on the development of this classification scheme can be found in “2013 NCHS Urban-Rural Classification Scheme for Counties” available at: [https://www.cdc.gov/nchs/data/series/sr\\_02/sr02\\_166.pdf](https://www.cdc.gov/nchs/data/series/sr_02/sr02_166.pdf).

---

## Sponsored Content

### *Food Related Programs*

The FOO section included a question sponsored by the USDA about participation in the Supplemental Nutrition Assistance Program (SNAP) in the last month. This question was asked of sample adult and sample child respondents who reported receiving SNAP benefits in the last 12 months. To reduce respondent burden, this question was asked once per family.

### *Family Food Security*

Family food security refers to access at all times to enough food for active, healthy lives. The food security questions in the FDS sections of the Sample Adult and Sample Child modules are sponsored by the USDA and consist of the same 10 questions measuring food security status of families in the past 30 days. The first three questions asked about being worried that food would not last until there is money to buy more, food bought did not last and didn't have money to buy more, and not being able to afford to eat balanced meals. The next set of statements asked about cutting or skipping meals, eating less than should, being hungry but not eating, losing weight, and not eating for a whole day because there wasn't enough money to buy food. Respondents were also asked for the number of days that a meal was skipped, and the number of days that did not eat for a whole day. During 2011-2018, the food security questions were administered as part of the Family Component at the beginning of the interview. Starting in 2019, the food security questions are administered once per family in the later portion of the Sample Adult and Sample Child modules using the family-level replicate interviewing approach. Responses to food security questions can be combined to create a raw food security score and categories for degree of food insecurity (see family food security recodes). The purpose of the questions is to examine the relationship between health and food insecurity. For more information about the USDA's food security research and standard procedures for measuring food insecurity and hunger in the United States, see <https://www.fns.usda.gov/sites/default/files/FSGuide.pdf>.

### *Family Food Security Recodes*

The ten questions in the Food Security (FDS) section were used to determine a 3 and 4 level scale of food security status of adults and children as recommended by the USDA Economic Research Service. The food security status classification variables were derived from a raw food security score created to represent the number of affirmative responses to the food security questions. Answers of "often true," "sometimes true," and "yes" are considered affirmative. Responses to questions that ask about the frequency of occurrence in the past 30 days are considered affirmative if the respondent's answer was greater than or equal to 3 days. Each affirmative response has a score of 1 for a total score ranging from 0 to 10. Not all ten questions were asked of all respondents. Respondents who answer "never true" to the first three questions (with a score of zero for each question) are determined to be food secured and are not asked additional questions. Subsequently, only those with an affirmative answer to questions that specify because there wasn't enough money for food, they cut the size or skipped meals, ate less than they should, were hungry but didn't eat, or lost weight, were asked about not eating for a whole day, and the number of days that occurred. Respondents who answered "don't know" or "refused" or whose answers were not ascertained to the first three questions in the food security set are classified as not ascertained food security status and coded as 8. Information from any affirmative response was summed to the raw score, including when respondents answered "don't know" or "refused" or whose answers were not ascertained for questions that came after the initial three food security questions within the set.

Two options for food security status classification variables were created: one with food security represented in a single “food secure” category, and one which distinguishes between families with high food security and families with marginal food security. The recommended classifications are given below:

Option 1

- Food secure (high or marginal food security, raw score 0–2)
- Low food security (raw score 3–5)
- Very low food security (raw score 6–10)

Option 2

- High food security (raw score 0)
- Marginal food security (raw score 1–2)
- Low food security (raw score 3–5)
- Very low food security (raw score 6–10)

---

## Telephone Usage

### Annual Core

The purpose of the cellular telephone questions is to track the use of wireless telephones in American families over time, allowing researchers to analyze the demographic characteristics of families who have substituted wireless service for landline home telephones. Having these data from a large population-based survey such as the NHIS provides useful information about potential bias from under coverage in random-digit-dial telephone surveys that use only landline telephone numbers in their sampling frames.

The redesigned questionnaire differs in several ways in the collection of telephone usage data from previous years and has the potential to impact estimates of telephone ownership for several reasons. First, prior to 2019, telephone ownership information was collected as part of the Family Core; a family respondent provided information for all persons living in responding families. Starting in 2019, the NHIS collects information for up-to-two persons living in each responding household, one randomly selected sample adult and one randomly selected sample child (if present in the household). The sample adult respondent may or may not also be the sample child respondent.

Second, the wireless telephone questions were modified, but the landline telephone question was not. To determine whether the sampled persons lived in a household with a landline telephone, the respondent was asked if there was “at least one phone inside your home that is currently working and is not a cell phone.” To avoid possible confusion with cordless landline telephones, the word “wireless” was not used in the survey. This question was asked only once, in whichever interview (Sample Adult or Sample Child) came first. In 2019, rather than asking family respondents whether “anyone in your family has a working cellular telephone,” the NHIS now asks sample adults whether they “have a working cell phone,” and if not, whether they “live with anyone who has a working cell phone.” This change permits the identification of adults living in wireless-only households (that is, households without landlines but with working cell phones) and of wireless-only adults (that is, adults who have a working cell phone and live in wireless-only households). Respondents for sample children are only asked the latter question, and only if the wireless status of the household is not yet known from the Sample Adult interview.

In 2019, an additional question was included for sample adults who have a cell phone and live in households with landline telephones. The Sample adult was asked to consider “all the telephone calls that you answer” and to report whether “all or almost all [are] on your cell phones, some [are] on your cell phone and some on your home phone, or very few or none [are] on your cell phones.” This question is a modified version of a question that, prior to 2019, was used to identify wireless-mostly households (rather than adults) based on calls received (rather than answered). The modified question permits the identification of “wireless-mostly” adults—defined as adults with both landline and cellular telephones who answer all or almost all calls on cell phones. Landline-mostly adults and dual-users can be similarly identified.

Third, the telephone ownership questions were moved from near the beginning of the survey to near the end of the survey. Breakoffs that occur late in the interview may result in interviews that are sufficiently complete for most purposes but are missing data for telephone ownership.

---

## NCHS Data Linkage Program

The Data Linkage Program at NCHS is a cross-cutting program housed in the Division of Analysis and Epidemiology (DAE) which aims to maximize the scientific value of the Center's population-based surveys, by linking NCHS survey data with data collected from vital and other administrative records. Linked data files enable researchers to augment information for major diseases, risk factors, and health service utilization, by linking exposures to outcomes and in some cases introducing a longitudinal component to survey data.

### Data Sources Linked

The Data Linkage Program currently links NHIS data with:

- National Death Index (NDI) death certificate data, including cause of death (<https://www.cdc.gov/nchs/data-linkage/mortality.htm>)
- Centers for Medicare and Medicaid Services (CMS) enrollment and claims data for:
  - o Medicare (<https://www.cdc.gov/nchs/data-linkage/medicare.htm>)
  - o Medicaid/CHIP (<https://www.cdc.gov/nchs/data-linkage/medicaid.htm>)
- Department of Housing and Urban Development's (HUD) administrative data from the largest rental housing assistance programs (<https://www.cdc.gov/nchs/data-linkage/hud.htm>)
- United States Renal Data System (USRDS) data on End Stage Renal Disease (ESRD) (<https://www.cdc.gov/nchs/data-linkage/esrd.htm>).
- Social Security Administration data, such as insurance claim data and utilization data, from Old Age, Survivors and Disability Insurance (OASDI) and Supplemental Security Income (SSI) benefit records (<https://www.cdc.gov/nchs/data-linkage/ssa.htm>).

For more information on available linked datasets, see <https://www.cdc.gov/nchs/data/datalinkage/LinkageTable.pdf>.

### Linkage Methodology

Only NHIS participants who have provided consent as well as the necessary personally identifiable information (PII) are considered *linkage-eligible*. *Linkage-eligibility* is distinct from program eligibility, which refers to whether a person meets eligibility criteria for a benefits program. *Linkage eligibility* refers to the potential ability to link data obtained from an NHIS participant to administrative data. Survey participants are informed of NCHS' intent to conduct data linkage activities through a variety of procedures such as "advance letters," participant brochures, and during the interview when verbal consent is requested. Starting in 2007, NHIS participants selected to be the sample adult or sample child (by proxy respondent) are asked for the last four digits of their Social Security Number (SSN) and Medicare Health Insurance Claim number (HICN) for participants age 65 and older. Additionally, those who refused to provide the last four digits of their SSN or HICN, are asked if they would consent to linkage based on their other identifying information. Only sample adult and sample child participants who provided the last four digits of SSN or HICN or provided consent for linkage without SSN or HICN are included in linkage activities for 2007 NHIS forward. Since 2010, approximately 90% of NHIS sample adult participants are linkage eligible. Questions to determine linkage eligibility are collected in the LNK sections of the Sample Adult and Sample Child interview.

The individual-level linkages are conducted using both probabilistic and deterministic techniques. The algorithms rely on PII such as SSN, name, and date of birth. Please refer to the appropriate linkage documentation for

further information on methodology and analytic considerations (for example for the linked NDI data, <https://www.cdc.gov/nchs/data-linkage/mortality-methods.htm>).

Sample addresses from NHIS are also geocoded to standard Census geocoded areas. This enables researchers to merge contextual data (e.g., air quality data) with NHIS data.

### Availability of Public-Use Linked Data

There are two types of public-use files released by the NCHS Data Linkage Program:

- a) Public-use linked mortality data files (<https://www.cdc.gov/nchs/data-linkage/mortality-public.htm>)
- b) Public-use feasibility data files released for the following data linkages:
  - o NCHS-CMS Medicare (<https://www.cdc.gov/nchs/data-linkage/medicare-feasibility.htm>)
  - o NCHS-CMS Medicaid (<https://www.cdc.gov/nchs/data-linkage/medicaid-feasibility.htm>)
  - o NCHS-SSA (<https://www.cdc.gov/nchs/data-linkage/ssa-feasibility.htm>)

Note: The feasibility files were developed to help interested researchers determine the maximum available sample sizes to assess the feasibility of conducting analyses utilizing the restricted-use linked files available through the NCHS Research Data Center.

### Restricted-Use Linked Data

All other linked data files are restricted-use and available only through the NCHS Research Data Center. For more information about the restricted-use linked data, including the file contents, methods used for linkage and analytic consideration, follow the links provided for each of the following data linkages:

- National Death Index (NDI), Restricted-Use Linked Mortality Data (<https://www.cdc.gov/nchs/data-linkage/mortality-restricted.htm>)
- NCHS-CMS Medicare (<https://www.cdc.gov/nchs/data-linkage/medicare-restricted.htm>)
- NCHS-CMS Medicaid (<https://www.cdc.gov/nchs/data-linkage/medicaid-restricted.htm>)
- NCHS-HUD (<https://www.cdc.gov/nchs/data-linkage/hud-restricted.htm>)
- NCHS- USRDS ESRD (<https://www.cdc.gov/nchs/data-linkage/esrd-restricted.htm>)
- NCHS-SSA (<https://www.cdc.gov/nchs/data-linkage/ssa-restricted.htm>)
- Geocoded data ([https://www.cdc.gov/rdc/geocodes/geowt\\_nhis.htm](https://www.cdc.gov/rdc/geocodes/geowt_nhis.htm))

For more information about accessing the restricted-use linked data, please visit the NCHS Research Data Center website: <https://www.cdc.gov/rdc/index.htm>



---

## Medical Expenditure Panel Survey (MEPS)

NHIS interviewed households also serve as a sampling frame for the Medical Expenditure Panel Survey (MEPS). MEPS, conducted by the Agency for Healthcare Research and Quality (AHRQ), collects data on the specific health services that Americans use, how frequently they use them, the cost of these services, and how they are paid for, as well as data on the cost, scope, and breadth of health insurance held by and available to U.S. workers.

The MEPS Household Component collects data from a nationally representative subsample of households that participated in the prior year's NHIS. Crosswalks that will allow data users to merge the MEPS full-year population characteristics public use data files with the NHIS person-level public use data files are available from AHRQ: [https://meps.ahrq.gov/mepsweb/data\\_stats/more\\_info\\_download\\_data\\_files.jsp#hc-nhis](https://meps.ahrq.gov/mepsweb/data_stats/more_info_download_data_files.jsp#hc-nhis).

---

## References

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

Bibbins-Domingo K; U.S. Preventive Services Task Force. (2016). Aspirin use for the primary prevention of cardiovascular disease and colorectal cancer: U.S. Preventive Services Task Force recommendation statement. *Annals of Internal Medicine*. 164:836–845.

Felitti, V. J., Anda, R. F., Nordenberg, D., Williamson, D. F., Spitz, A. M., Edwards, V., Koss, M. P., & Marks, J. S. (1998). Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults: The Adverse Childhood Experiences (ACE) study. *American Journal of Preventive Medicine*. 14:245-258.

Goodman A, Goodman R. (2009). Strengths and difficulties questionnaire as a dimensional measure of child mental health. *Journal of the American Academy of Child and Adolescent Psychiatry*.48:400-3.

Goodman R. (1997). The Strengths and Difficulties Questionnaire: a research note. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*. 38(5):581-6.

Ingram DD, Franco SJ. (2014). 2013 NCHS urban–rural classification scheme for counties. *National Center for Health Statistics. Vital Health Stat* 2(166).

Korn EL, Graubard BI. (1999). *Analysis of Health Surveys*, John Wiley & Sons.

Kroenke K, Spitzer RL, Williams JBW. (2001). The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine*. 16: 606–613.

Kroenke K, Spitzer RL, Williams JB. (2003). The Patient Health Questionnaire-2: Validity of a Two-Item Depression Screener. *Medical Care*. 41:1284-92.

Kroenke K, Spitzer RL, Williams JB, Monahan PO, Löwe B. (2007). Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. *Annals of Internal Medicine*. 146(5):317-325.

Kroenke K, Strine TW, Spitzer RL, Williams JB, et al. (2009) The PHQ-8 as a measure of current depression in the general population. *Journal of Affective Disorders* 114, 163–173.

Kroenke K, Spitzer RL, Williams JB, Löwe B. (2010). The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: a systematic review. *General Hospital Psychiatry*. 32(4):345-359.

Löwe B, Decker O, Müller S, Brähler E, et al. (2008). Validation and standardization of the generalized anxiety disorder screener (GAD-7) in the general population. *Medical Care*. 46, 266–274.

Olson K. (2013). Paradata for nonresponse adjustment. *The Annals of the American Academy of Political and Social Science* 645(1):142-170.

Qaseem A, Crandall CJ, Mustafa RA, Hicks LA, Wilt TJ; Clinical Guidelines Committee of the American College of Physicians. (2019). Screening for Colorectal Cancer in Asymptomatic Average-Risk Adults: A Guidance Statement from the American College of Physicians. *Annals of Internal Medicine*. 171(9):643-654.

- 
- Office of Management and Budget. (1997). Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity. Federal Register 62(210):58782-58790.
- Rutter LA, Brown TA. (2017). Psychometric Properties of the Generalized Anxiety Disorder Scale-7 (GAD-7) in Outpatients with Anxiety and Mood Disorders. *Journal of Psychopathology and Behavioral Assessment*. 39(1):140-146.
- Spitzer RL, Kroenke K, Williams JB, Löwe B. (2006). Brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine*. 166:1092-7.
- Sheldrick RC, Henson BS, Neger EN, Merchant S, Murphy JM, Perrin EC. (2013). The Baby Pediatric Symptom Checklist: Development and initial validation of a new social/emotional screening instrument for very young children. *Academic Pediatrics* 13:72-80.
- Stussman BJ, Taylor BL, Riddick H. (2003). Partials and Break-offs in the National Health Interview Survey, 2002. Paper presented at the 2003 Federal Committee on Statistical Methodology Research Conference, Arlington, Virginia.
- Valliant R, Dever JA, Kreuter F. (2018). *Practical tools for designing and weighting survey samples*. New York: Springer.
- van Buuren S. (2012). *Flexible Imputation of Missing Data (2nd Edition)*. Boca Raton, FL: Chapman and Hall/CRC.

## Appendix: Availability of Health, Health Insurance and Selected Demographic Questions Not Included in the Sample Adult and Sample Child Public Use Files

Table. List of questions not available in the public use files and availability through the Research Data Center (D), as a recode (R), or not available (NA): 2019 NHIS

Module	Section Acronym	Questionnaire Variable	Description	Availability	Public Use Recode
Roster	HHC	EDUC	Highest level of education completed	D R	EDUC_A, MAXEDUC_A, PRTNREDUC_A, SPOUSEDUC_A, MAXEDUC_C, MAXPAREduc_C
Roster, Sample Adult, Sample Child	HHC, VFY	RACE, RACE_SP, RACE_VRBAT, NEWRACE_A, NATORG, NEWNATORG_A, HISPTYPE_A, HISPOTHER_A, HISPVRBAT_A, PITYPE_A, PIOTHER_A, PIVRBAT_A, ASIANTYPE_A, ASIANOTHER_A, ASIANVRBAT_A, RACEOTHER_A, RACEVRBAT_A, MLTRACE_A, NEWNATORG_C, HISPTYPE_C, HISPOTHER_C, HISPVRBAT_C, PITYPE_C, PIOTHER_C, PIVRBAT_C, ASIANTYPE_C, ASIANOTHER_C, ASIANVRBAT_C, RACEOTHER_C, RACEVRBAT_C, MLTRACE_C	Race and Ethnicity	D R	RACEALLP_A, HISPALLP_A, HISP_A, HISDETP_A, RACEALLP_C, HISPALLP_C, HISP_C, HISDETP_C, SASPPRACE_A, SASPPHISP_A, SCPARRAC_C, RACPARSC_C, SCPARHISP_C, HISPPARSC_C
Sample Adult	BMI	HEIGHTFT_A, HEIGHTIN_A, HEIGHTM_A, HEIGHTCM_A	Height	D R	HEIGHTTC_A
Sample Adult	BMI	WEIGHTLB_A, WEIGHTKG_A	Weight	D R	WEIGHTLBT_C
Sample Adult	CAN	CANKIND1_A-CANKIND3_A	Kind of cancer	D R	BLADDCAN_A BLOODCAN_A BONECAN_A BRAINCAN_A BREASCAN_A CERVICAN_A COLONCAN_A ESOPHCAN_A GALLBCAN_A LARYNCAN_A LEUKECAN_ALIVERCAN_A LUNGCAN_A LYMPHCAN_A MELANCAN_A MOUTHCAN_A OVARYCAN_A PANCRCAN_A PROSTCAN_A RECTUCAN_A SKNMCAN_A SKNNMCAN_A SKNDKCAN_A STOMACAN_A THROCAN_A THYROCAN_A UTERUCAN_A HDNCKCAN_A COLRCCAN_A OTHERCAN_A
Sample Adult	CAN	CANKIND1_A-CANKIND3_A	Kind of cancer (kidney, testicular)	D	
Sample Adult	CAN	CANAGE1_A-CANAGE3_A	Age of cancer diagnosis	D R	BLADDAGETC_A, BLOODAGETC_A, BONEAGETC_A, BRAINAGETC_A, BREASAGETC_A, CERVIAGETC_A,

					COLONAGETC_A, ESOPHAGETC_A, GALLBAGETC_A, LARYNAGETC_A, LEUKEAGETC_A, LIVERAGETC_A, LUNGAGETC_A, LYMPHAGETC_A, MELANAGETC_A, MOUTHAGETC_A, OVARYAGETC_A, PANCRAGETC_A, PROSTAGETC_A, RECTUAGETC_A, SKNMAGETC_A, SKNNMAGETC_A, SKNDKAGETC_A, STOMAAGETC_A, THROAAGETC_A, THYROAGETC_A, UTERUAGETC_A, HDNCKAGETC_A, COLRCAGETC_A, OTHERAGETC_A
Sample Adult	CAN	CANAGE1_A-CANAGE3_A	Age of cancer diagnosis (kidney, testicular)	D	
Sample Adult	CAN	CANMORE_A	More than three kinds of cancer	D R	NUMCAN_A
Sample Adult	DIB	DIBAGE_A	Age of diabetes diagnosis	D R	DIBAGETC_A, DIFYRSTC_A
Sample Adult	EMP	EMPLASTWK_A	Working last week	D R	EMPWRKLSWK_A
Sample Adult	EMP	EMPNOWRK_A	Temporarily absent from work last week	D R	EMPWRKLSWK_A
Sample Adult	EMP	EMPWRKHRS_A	Number of hours worked last week	D R	EMPWKHRS2_A
Sample Adult	EMP	EMPWKFT_A	Usually working 35+ hours per week	D R	EMPWRKFT_A
Sample Adult	EMP	EMPLSTWRK_A	Last time worked for pay	D R	EMPLSTWORK_A
Sample Adult	EMP	EMPDYSMSS_A	Days missed work in past 12 months due to illness/injury/disability	D R	EMPDYSMSS2_A
Sample Adult	FEM	FEMWORK_A	Employment status of other adults	D R	PCNTADTWRP_A
Sample Adult	FEM	FEMWKFT_A	Other adult family member works 35+ hours per week	D R	PCNTADTWFP_A
Sample Child	FEM	FEMWORK_C	Employment status of adults	D R	PCNTADTWRP_C, PCNTPARWKP_C
Sample Child	FEM	FEMWKFT_C	Adult family member works 35+ hours per week	D R	PCNTADTWFP_C, PCNTPARWFP_C
Sample Adult	IMS	ZOSTAVAXYR_A	Year of most recent Zostavax vaccine	D R	ZOSTAVAXYRP_A
Sample Adult	IMS	SHINGRIXYR_A	Year of most recent Shingrix vaccine	D R	SHINGRIXYRP_A
Sample Adult	IMS	SHTHPVAGE_A	Age of first HPV vaccine	D R	SHTHPVAGEP_A
Sample Adult	INC	INCTOTAL_A	Total family income	D R	FAMINCTC_A, INCGRP1_A, POVRATTC_A, RATCAT_A
Sample Adult	INC	INC400PCT_A, INC250PCT_A, INC200PCT_A, INC138PCT_A, INC100PCT_A	Family income poverty levels	D	
Sample Adult	INC	INC150K_A, INC100K_A, INC75K_A	Family income ranges	D	
Sample Child	INC	INCTOTAL_C	Total family income	D R	FAMINCTC_C, INCGRP1_C, POVRATTC_C, RATCAT_C
Sample Child	INC	INC400PCT_C, INC250PCT_C, INC200PCT_C, INC138PCT_C, INC100PCT_C	Family income poverty levels	D	
Sample Child	INC	INC150K_C, INC100K_C, INC75K_C	Family income ranges	D	
Sample Adult	INS	HIKIND_A	Kinds of health insurance	R	MEDICARE_A, MEDICAID_A, PRIVATE_A, CHIP_A, OTHPUB_A, OTHGOV_A, MILITARY_A, IHS_A, NOTCOV_A, COVER_A, COVER65_A,

					HIKIND01_A, HIKIND02_A, HIKIND03_A, HIKIND04_A, HIKIND05_A, HIKIND06_A, HIKIND07_A, HIKIND08_A, HIKIND09_A, HIKIND10_A
Sample Adult	INS	HICHANGE_A	Verification of insurance coverage	NA	
Sample Adult	INS	MCANAME_A	Verbatim response to name of Medicare Advantage or Medicare HMO plan	R	MCADVR_A
Sample Adult	INS	MACHMN_A	Verbatim response to name of Medicaid managed care plan	NA	
Sample Adult	INS	PLANNAME1_A	Adult shares child's plan 1	R	PRIVATE_A
Sample Adult	INS	POLHLDA1_A	Policyholder for adult who shares child's plan 1	R	POLHLD1_A, POLHLD2_A, PRPLCOV1_A, PRPLCOV2_A, PRPOLH1_A, PRPOLH2_A
Sample Adult	INS	PRPOLHP1_A	Relationship to policyholder for adult who shares child's plan 1	R	POLHLD1_A, POLHLD2_A, PRPLCOV1_A, PRPLCOV2_A, PRPOLH1_A, PRPOLH2_A
Sample Adult	INS	PLANNAME2_A	Adult shares child's plan 2	R	PRIVATE_A
Sample Adult	INS	POLHLDA2_A	Policyholder for adult who shares child's plan 2	R	POLHLD1_A, POLHLD2_A, PRPLCOV1_A, PRPLCOV2_A, PRPOLH1_A, PRPOLH2_A
Sample Adult	INS	PRPOLHP2_A	Relationship to policyholder for adult who shares child's plan 2	R	POLHLD1_A, POLHLD2_A, PRPLCOV1_A, PRPLCOV2_A, PRPOLH1_A, PRPOLH2_A
Sample Adult	INS	HIPNAM1_A	Verbatim response to name of Sample Adult's first private plan	R	EXCHPR1_A
Sample Adult	INS	MORPLAN_A	Any other plans	NA	
Sample Adult	INS	HIPNAM2_A	Verbatim response to name of Sample Adult's second private plan	R	EXCHPR2_A
Sample Adult	INS	POLHLD_A	Policyholder for private plan	R	POLHLD1_A, POLHLD2_A
Sample Adult	INS	PRPLCOV_A	Plan cover others	R	PRPLCOV1_A, PRPLCOV2_A
Sample Adult	INS	PRPOLH_A	Relationship to policyholder	R	PRPOLH1_A, PRPOLH2_A
Sample Adult	INS	PLNWRK_A	How adult's plan was obtained	R	PLNWRKR1_A, PLNWRKR2_A
Sample Adult	INS	PLNWKSP_A	Verbatim response to how plan was obtained	R	PLNWRKR1_A, PLNWRKR2_A
Sample Adult	INS	PLNEXCHG_A	Plan obtained through the Marketplace	R	PLNEXCHG2_A, PLNEXCHG1_A
Sample Adult	INS	PLNPAY_A	Who pays for this plan	R	PLN1PAY1_A, PLN1PAY2_A, PLN1PAY3_A, PLN1PAY4_A, PLN1PAY5_A, PLN1PAY6_A, PLN2PAY1_A, PLN2PAY2_A, PLN2PAY3_A, PLN2PAY4_A, PLN2PAY5_A, PLN2PAY6_A
Sample Adult	INS	HICOSTN_A, HICOSTT_A	Premium amount that family or adult pays for plan	R	HICOSTR1_A, HICOSTR2_A
Sample Adult	INS	PRDEDUC_A	Plan has a deductible	R	PRDEDUC1_A, PRDEDUC2_A
Sample Adult	INS	PRHDHP_A	Annual deductible	R	PRHDHP1_A, PRHDHP2_A
Sample Adult	INS	HSAHRA_A	Health savings account	R	HSAHRA1_A, HSAHRA2_A
Sample Adult	INS	PRRXCOV_A	Plan has prescription drug coverage	R	PRRXCOV1_A, PRPXCOV2_A
Sample Adult	INS	PRDNCOV_A	Plan has dental coverage	R	PRDNCOV1_A, PRDNCOV2_A
Sample Adult	INS	PRVSCOV_A	Plan has vision coverage	R	PRVSCOV1_A, PRVSCOV2_A

Sample Adult	INS	CHNAME_A	Verbatim response to name of Sample Adults' Children's Health Insurance Program (CHIP) plan	NA	
Sample Adult	INS	OPNAME_A	Verbatim response to name of Sample Adults' state-sponsored plan	R	PLEXCHOP_A
Sample Adult	INS	OGNAME_A	Verbatim response to name of Sample Adults' other government plan	R	PLEXCHOGR_A
Sample Adult	INS	MILSPC_A	Type of military related health care	R	MILSPC1_A, MILSPC1R_A, MILSPC2_A, MILSPC3_A
Sample Adult	INS	RSNHIOHSP_A	Verbatim response to reasons for not getting coverage	R	RSNHICOST_A, RSNHIWANT_A, RSNHIELIG_A, RSNHICONF_A, RSNHIMEET_A, RSNHIWAIT_A, RSNHIOH_A, RSNHIJOB_A, RSNHIMISS_A
Sample Child	INS	HIKIND_C	Kinds of health insurance	R	MEDICAID_C, PRIVATE_C, CHIP_C, OTHPUB_C, OTHGOVR_C, MILITARY_C, IHS_C, NOTCOV_C, COVER_C, HIKIND01_C, HIKIND03_C, HIKIND04_C, HIKIND05_C, HIKIND06_C, HIKIND07_C, HIKIND08_C, HIKIND09R_C, HIKIND10_C
Sample Child	INS	HICHANGE_C	Verification of insurance coverage	NA	
Sample Child	INS	MCANAME_C	Verbatim response to name of Medicare Advantage or Medicare HMO plan	D, R	
Sample Child	INS	MCPART_C	Type of Medicare coverage	D	
Sample Child	INS	MCCHOICE_C	Enrolled in Medicare Advantage Plan	D	
Sample Child	INS	MCHMO_C	Medicare HMO	D	
Sample Child	INS	MCPARTD_C	Medicare Part D	D	
Sample Child	INS	MACHMN_C	Verbatim response to name of Medicaid managed care plan	NA	
Sample Child	INS	PLANNAME1_C	Child shares adults' plan 1	R	PRIVATE_C
Sample Child	INS	POLHLD1_C	Policyholder for child who shares adult's plan 1	R	POLHLD1_C, POLHLD2_C, PRPLCOV1_C, PRPLCOV2_C, PRPOLH1_C, PRPOLH2_C
Sample Child	INS	PLANNAME2_C	Child shares adults' plan 2	R	PRIVATE_C
Sample Child	INS	POLHLD2_C	Policyholder for child who shares adult's plan 2	R	POLHLD1_C, POLHLD2_C, PRPLCOV1_C, PRPLCOV2_C, PRPOLH1_C, PRPOLH2_C
Sample Child	INS	HIPNAM1_C	Verbatim response to name of Sample Child's first private plan	R	EXCHPR1_C
Sample Child	INS	MORPLAN_C	Any other plans	NA	
Sample Child	INS	HIPNAM2_C	Verbatim response to name of Sample Child's second private plan	R	EXCHPR2_C
Sample Child	INS	POLHLD_C	Policyholder for private plan	R	POLHLD1_C, POLHLD2_C
Sample Child	INS	PRPLCOV_C	Plan cover others	R	PRPLCOV1_C, PRPLCOV2_C
Sample Child	INS	PLNWRK_C	How plan was obtained	R	PLNWRKR1_C, PLNWRKR2_C
Sample Child	INS	PLNWKSP_C	Verbatim response to how plan was obtained	R	PLNWRKR1_C, PLNWRKR2_C
Sample Child	INS	PLNEXCHG_C	Plan obtained through the Marketplace	R	PLNEXCHG2_C, PLNEXCHG1_C
Sample Child	INS	PLNPAY_C	Who pays for this plan	D, R	PLN1PAY1_C, PLN1PAY2_C, PLN1PAY3_C, PLN1PAY5_C, PLN1PAY6R_C, PLN2PAY1_C, PLN2PAY2_C, PLN2PAY3_C, PLN2PAY5_C, PLN2PAY6R_C

Sample Child	INS	HICOSTN_C, HICOSTT_C	Premium amount that family pays for plan	R	HICOSTR1_C, HICOSTR2_C
Sample Child	INS	PRDEDUC_C	Plan has a deductible	R	PRDEDUC1_C, PRDEDUC2_C
Sample Child	INS	PRHDHP_C	Annual deductible	R	PRHDHP1_C, PRHDHP2_C
Sample Child	INS	HSAHRA_C	Health savings account	R	HSAHRA1_C, HSAHRA2_C
Sample Child	INS	PRRXCOV_C	Plan has prescription drug coverage	R	PRRXCOV1_C, PRPXCOV2_C
Sample Child	INS	PRDNCOV_C	Plan has dental coverage	R	PRDNCOV1_C, PRDNCOV2_C
Sample Child	INS	PRVSCOV_C	Plan has vision coverage	R	PRVSCOV1_C, PRVSCOV2_C
Sample Child	INS	CHNAME_C	Verbatim response to name of Sample Child's Children's Health Insurance Program (CHIP) plan	NA	
Sample Adult	MAR	SPOUSWHO_A	Person number for spouse	D	
Sample Adult	MAR	SPOUSSEX_A, SPOUNWSEX_A	Confirm spouse's sex; Correct spouse's sex	D R	SPOUSESEX_A
Sample Adult	MAR	PARTNERWHO_A	Person number for partner	D	
Sample Adult	MAR	PARTNERSEX_A, PARTNEWSEX_A	Confirm partner's sex; Correct partner's sex	D R	PRTNRSEX_A
Sample Adult	MAR	LEGALSTAT_A	Legal marital status	D	LEGMARSTAT_A
Sample Adult	MAR	WIDIVSEP_A	Widowed/Divorced/Separated	D R	MARSTAT_A
Sample Adult	NAT	CITIZEN_A	Citizen status	D R	CITZNSTP_A
Sample Adult	NAT	NATSTBORN_A	Place of birth	D	
Sample Adult	NAT	NATCTZN_A	How sample adult became U.S. citizen	D	
Sample Child	NAT	CITIZEN_C	Citizen status	D R	CITZNSTP_C
Sample Child	NAT	NATSTBORN_C	Place of birth	D	
Sample Child	NAT	NATCTZN_C	How sample child became U.S. citizen	D	
Sample Child	PAR	RELCHPAR_C	Biological/Adoptive/Step/Other to type of parent-child relationship	D R	RELCHPARP1_C-RELCHPARP2_C
Sample Child	PAR	MARITAL_C	Married/Living with partner as unmarried couple/neither to questions about marital status of sample child's parents	D R	MARITAL1_C-MARITAL2_C
Sample Child	PAR	SPOUSLIV_C	Spouse of first-fourth parent lives there	D	
Sample Child	PAR	SPOUSEP_C	First-fourth parents are legally separated	D	
Sample Child	PAR	SPOUSWHO_C	Person number of first-fourth residential parent's spouse	D	
Sample Child	PAR	SPOUSSEX_C; FIXSPOUSSEX_C_C	Confirming sex of first-fourth parent's spouse; Correcting sex of first-fourth parent's spouse	D	
Sample Child	PAR	PARTNERWHO_C	Person number of first-fourth residential parent's unmarried partner	D	
Sample Child	PAR	PARTNERSEX_C; FIXPARTSEX_C	Confirming sex of first-fourth parent's unmarried partner; Correcting sex of first-fourth parent's unmarried partner	D	
Sample Child	PAR	EVRMARRIED_C, WIDIVSEP_C	Ever been married	D R	MARSTAT1_C-MARSTAT2_C
Sample Child	PAR	LEGALSTAT_C	Married/Widowed/Divorced/Separated to questions about legal marital status	D R	LEGMSTAT1_C-LEGMSTAT2_C
Sample Child	PAR	PARBORN_C	Sample child's first-second parent born in the US/US territory	D	



Sample Child	PAR	FOSTPAR	Sample Child currently in foster care	D	
Sample Adult	PRV	PSA5YR_A	Number of PSA tests in last 5 years	D R	PSA5YRTC_A
Sample Adult	PRV	COLKIND_A	Tests to check for colon cancer	R	COLKIND1_A- COLKIND6_A
Sample Adult	PRV	CGUARDWHEN_A	Most recent Cologuard test	NA	
Sample Adult	SCH	SCHDYSMSS_A	Number of school days missed, past 12m	D R	SCHDYSSTC_A
Sample Child	SCH	SCHDYSMSS_C	Number of school days missed, past 12m	D R	SCHDYSSTC_C
Sample Adult	UTZ	EMERGE12M_A	Number of times visited hospital emergency room, past 12m	D R	EMERG12MTC_A
Sample Adult	UTZ	URGENT12M_A	Number of times visited urgent care, past 12m	D R	URGNT12MTC_A
Sample Child	UTZ	EMERGE12M_C	Number of times visited hospital emergency room, past 12m	D R	EMERG12MTC_C
Sample Child	UTZ	URGENT12M_C	Number of times visited urgent care, past 12m	D R	URGNT12MTC_C

NOTE: The Research Data Center (RDC) is a data enclave established to provide a mechanism whereby researchers can access detailed data files in a secure environment without jeopardizing the confidentiality of survey participants. Information about RDC access options and application procedures is available at: <https://www.cdc.gov/rdc/>.