2021
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

August 2022

Table of Contents

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 2021 NHIS and previous years, as well as important information about any modifications or updates to the data 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 mailing list. The Listserv is made up of over 3,150 NHIS data users located around the world who receive 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, and select "National Health Interview Survey (NHIS) Researchers" as one of your options, or click on "Contact Us" on the NHIS website and scroll to "Listserv" or "How to Subscribe to the HISUSERS List."

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. 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 2021 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, 2021

In a bibliography, the suggested citation for this document is:

> National Center for Health Statistics. National Health Interview Survey, 2021 survey description. 2022. Available from: https://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/NHIS/2021/srvydesc-508.pdf

The suggested citation for 2021 NHIS survey data and other documentation is:

> National Center for Health Statistics. National Health Interview Survey, 2021. Public-use data file and documentation. https://www.cdc.gov/nchs/nhis/data-questionnaires-documentation.htm. 2022.

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 NCHS publications and 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 2021?

- Due to ongoing data collection difficulties posed by the COVID-19 pandemic, NHIS cases continued to be attempted by telephone first from January to April 2021. Personal visits were used only to follow-up on nonresponse, deliver recruitment materials, and conduct interviews when telephone numbers were unknown. Starting in May 2021, interviewers were instructed to return to regular survey interviewing procedures, whereby first contact attempts to households were made in person, with follow-up allowed by telephone. Interviewers were given flexibility to continue using telephone first contact attempts based on local COVID-19 conditions.

- Some modifications were made to the weighting process for 2021 such as the use of Recursive Partitioning for Modeling Survey Data for estimation of complex interactions, and the use of substitute calibration totals from the 2021 Current Population Survey in the absence of single-year American Community Survey estimates for population characteristics of interest.

- Beginning in 2021, the annual content of the employment section fielded during 2019-2020 was restructured to ask for the reason for not working for pay last week and the last time the Sample Adult worked for pay earlier in the set of questions. This change informed the universe of subsequent questions. As a result, variables names related to employment, industry and occupation, and social distancing at work have been updated.

- In addition to annual content, 2021 rotating, sponsored and emerging content is listed below:

Sample Adult Rotating Core

January-December:

Allergies, weak or failing kidneys, hepatitis, cirrhosis or other long-term liver conditions, hearing in a noisy and quiet room, use of sign language, chronic pain and pain location, preventive screening related to colon, breast, cervical, and prostate cancer, checks for blood pressure, cholesterol, and glucose, serious psychological distress, injuries (sudden and repetitive strain injuries), and industry and occupation,

April-December:

Aspirin use

Sample Child Rotating Core

January-December:

Allergies, mental health care, injuries, stressful life events

Sample Adult Sponsored Content

January-December:

Life satisfaction, distress due to living with diabetes, insulin use, AIC measures, chronic fatigue syndrome, prolonged cold or flu, dry mouth, epilepsy, problems with taste and smell, COVID-19 symptoms related to taste and smell, age of disability onset, immunization (flu, pneumonia, shingles, tetanus, and Hepatitis A/B vaccine), cancer treatment during coronavirus pandemic, detailed preventive screening questions related to colon, breast, cervical, and prostate cancer, noncigarette tobacco product use, work arrangements, social distancing at work, food security, and food related programs

April-December:

COVID-19 vaccination

Sample Child Sponsored Content

January-December:

Life satisfaction, additional measures of stressful life events, food security, and food related programs

July-December:

COVID-19 vaccination among 12–17-year-olds

Sample Adult Emerging Content

January-December:

Immunosuppression, access to care during the pandemic, caregiving received, telehealth, social and emotional support, and COVID-19 diagnosis

Sample Child Emerging Content

January-December:

Interaction with health care providers, access to care during the coronavirus pandemic, telehealth, social and emotional support, additional measures for stressful life events, bullying, lifetime concussions, and COVID-19 diagnosis

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.

Overview of 2021 Survey Methods

NHIS is a nationally representative household survey of the U.S. civilian noninstitutionalized population. It is conducted continuously throughout the year by the National Center for Health Statistics (NCHS). Interviews are typically conducted in respondents' homes, but follow-ups to complete interviews may be conducted over the telephone.

Due to the COVID-19 pandemic, typical data collection procedures were disrupted. During January through April 2021, contact with household members was attempted first via telephone, with subsequent personal visits allowed (a continuation of procedures from the latter part of 2020). Beginning in May 2021, interviewers returned to regular survey interviewing procedures, whereby first contact with household members was attempted in person, with follow-up allowed by telephone. Allowed personal visits varied by Regional Office, based on COVID-19 rates in local jurisdictions. Interviewers were required to wear masks, practice social distancing, and be fully vaccinated if interviews were conducted in the home. In 2021, 62.8% of the Sample Adult interviews and 61.4% of the Sample Child interviews were conducted at least partially by telephone.

Information about the Sample Adult is self-reported, unless physically or mentally unable to do so and a knowledgeable proxy can answer for the Sample Adult. Information about the Sample Child is collected from a parent or adult who is knowledgeable about and responsible for the health care of the Sample Child. In 2021, there were 29,482 Sample Adult interviews and 8,261 Sample Child interviews. The Sample Adult response rate was 50.9% and the Sample Child response rate was 49.9%. The NHIS includes annual content that appears on the survey every year. The survey also includes rotating core content, sponsored content, and emerging content that appears periodically. Visit https://www.cdc.gov/nchs/nhis/2019_quest_redesign.htm for a description of content in any given year. For more information about NHIS, visit: https://www.cdc.gov/nchs/nhis.htm.

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 are not 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 typically 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 2016, 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). At that time 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. For more information about the 2016–2025 NHIS sample design and estimation structures, see Vital and Health Statistics, Series 2, Number 191 (cdc.gov)

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. In 2021, about 717 interviewers, also called "Field Representatives" or "FRs," conducted NHIS interviews nationwide. FRs are trained and directed by health survey supervisors in the U.S. Census Bureau Regional Offices. 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.

Typically, 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.

Interviewing During the COVID Pandemic

The COVID-19 pandemic continued to impact NHIS field work procedures in 2021 relating to first attempt —the first attempt that an FR makes, in-person or by phone, to determine eligibility of a household at a sample address and recruit a household participant in the survey, regardless of whether contact was made; and first contact—the first interaction between the FR and anyone from the household at a sample address, in-person or by phone.

During January through April, first contact with households was attempted first via telephone, with subsequent personal visits allowed (a continuation of procedures from the latter part of 2020). Allowed personal visits varied by Regional Office, based on COVID-19 rates in local jurisdictions. Regional Offices' mailing with the Advance letter to selected sample addresses also mentioned limited personal visits from field staff. While FRs were instructed to maximize phone attempts, circumstances for personal visits included: no working phone numbers; no answer via phone; confirmation of units are either unoccupied or occupied solely by persons not eligible for interview (type B) or units no longer qualify as a housing unit (Type C); locate addresses; drop off survey materials; and follow-up on non-response cases. In addition, when FR made contact at the door, they were instructed to attempt to schedule a phone interview before conducting the interview from a porch or hallway. Interviewers in the field were required to wear masks (covering both the nose and the mouth), practice social distancing by staying six feet away, no direct hand to hand contact (e.g., avoid handing out a brochure), and practicing hand hygiene (i.e., handwashing or use of hand sanitizer).

To reach households by telephone, interviewers use the telephone numbers appended with the address-based sample from commercial address lists. For NHIS sample addresses without a telephone match, FRs used online phone number searches to identify household telephone numbers.

U.S. Census Bureau issued guidance to FRs for telephone interviewing.

<u>When speaking to a person:</u>

- Identify themselves, affiliation, and purpose of call
  "Hello, my name is (say your name), I'm calling from the U.S. Census Bureau, on behalf of the National Center for Health Statistics, which is part of the Centers for Disease Control and Prevention. We are conducting a health survey. This is a nationwide survey about the health of both adults and children."*
- For privacy concerns, only mention the name of the survey after confirming to be speaking to someone at the correct address
- For safety concerns, confirm that the person is not driving. If driving, call another time
- Mention the "Advance letter" and ask if the household has received the letter
- Confirm that the person is over 18
- Confirm the correct sample address and that the person usually lives at sample address
- Provide instructions to the household respondent on how to verify their identity and employment status with Census

<u>When leaving a voicemail message:</u>

- Identify themselves and affiliation, mention the "Advance letter," purpose of call, and change from in-person visit to telephone call (when applicable), telephone number for returning call, instructions for verifying their employment with Census, repeat their name and contact telephone number, mention confidentiality of survey information, and a conclude greeting.
- Not mention the name of the survey in the voice message

<u>When non-contact or refusals:</u>

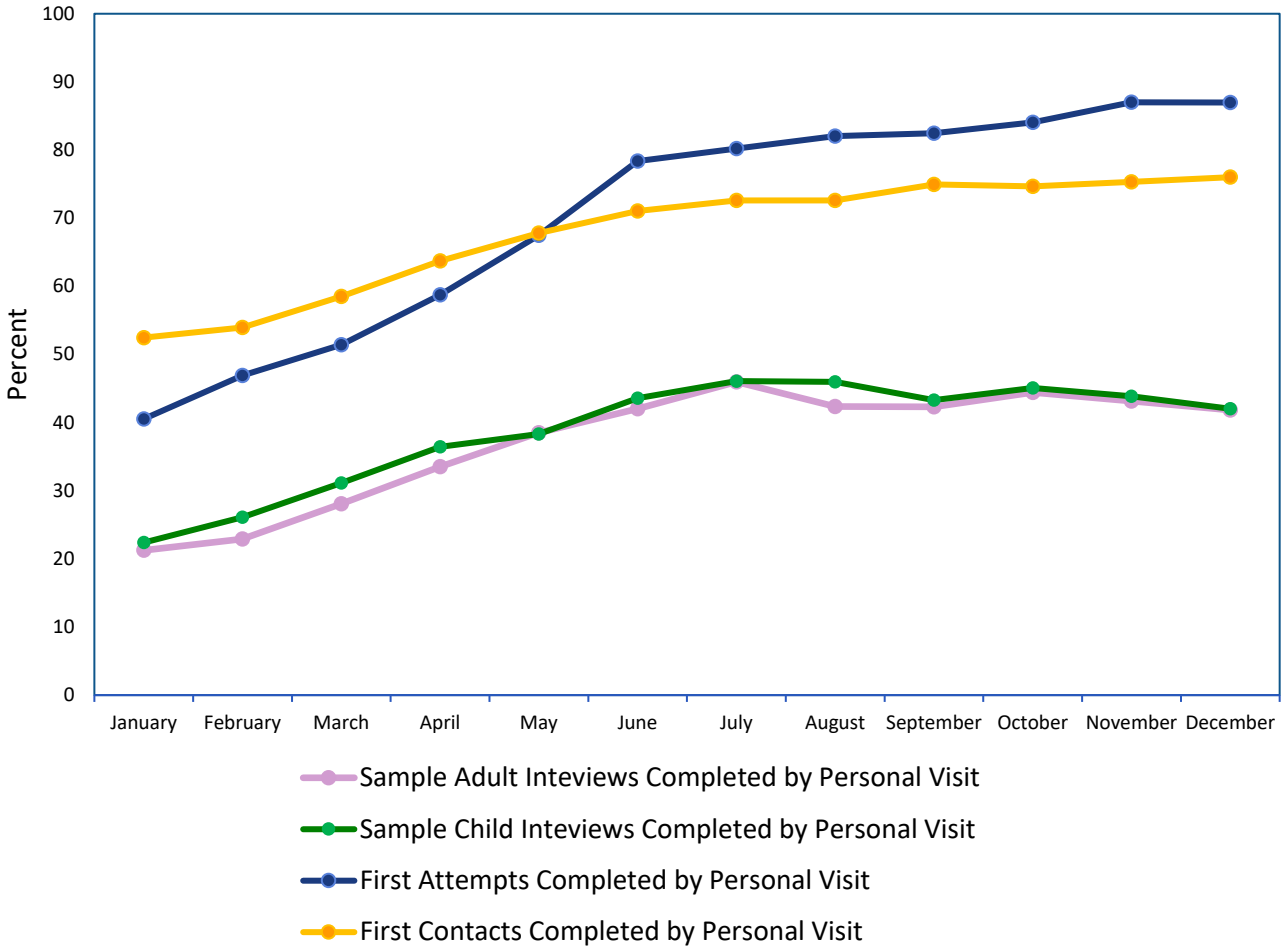
- Limits on number of calls and spacing between calls made

In May 2021, interviewers returned to regular survey interviewing procedures, whereby first attempts to households were made in person, with follow-up allowed by telephone. Interviewers continued to be required to wear masks, have no direct hand to hand contact, and practice social distancing in the field, and were required to be fully vaccinated (defined as two weeks after the second dose in a two-dose series, or a single-dose vaccine) if interviewers were conducted in the home.

Data Collection and Field Results

In 2021, 62.8% of the Sample Adult interviews and 61.4% of the Sample Child interviews were conducted at least partially by telephone. This is lower than the percentage of telephone interviews that were conducted at least partially by telephone in 2020 (70.7% for Sample Adult interviews and 68.0% for Sample Child interviews), but still higher than interviews in 2019 for both the Sample Adult (34.3%) and Sample Child (31.7%).

Figure 1 shows metrics summarizing the extent to which the NHIS returned to personal visits throughout the calendar year including the percent of Sample Adult and Sample Child interviews completed by personal visit, the percent of first attempts made by personal visit, and the percent of first contacts completed by personal visit. The figure shows that the percentage of both Sample Adult and Sample Child interviews completed by personal visit increased from about 21% in January to 46% in July, and then mostly leveled off for the remainder of the year. By its complement, Sample Adult and Sample Child interviews completed by phone was about 79% in January, decreased to 54% in July and remained over fifty percent during July-December. First attempt and first contact by personal visit increased sharply from January to June, followed with a gradual increase thereafter. About 41% of first attempts were made by personal visit in January, compared to 87% in November and December. First contact by personal visit increased from 53% in January to 76% in December.



Rostering and Respondents

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 aged 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 2021, 93.8% of the Sample Child respondents were the child's parent, either a biological, adoptive or stepparent, while 3.9% were a grandparent, 0.5% were an aunt or uncle, 0.7% were an adult sibling, 1.0% were another relative or other non-relative, and 0.07% were not ascertained. 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.

Confidentiality

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 Field Representative Manual available on the NHIS website, https://www.cdc.gov/nchs/nhis.htm.

III. NHIS Redesign

Beginning in 2019, the structure and content of the NHIS questionnaire was 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

Structure of 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 themself unless they are 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.

In the redesigned survey, family-level content is collected in the Sample Adult and/or Sample Child questionnaire module. Figure 3 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. 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 illustrates the 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 2. NHIS Redesigned Structure

Households with at least one adult and one child

or

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.

Households with at least one adult and no children

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.

Content of the Redesigned Questionnaire

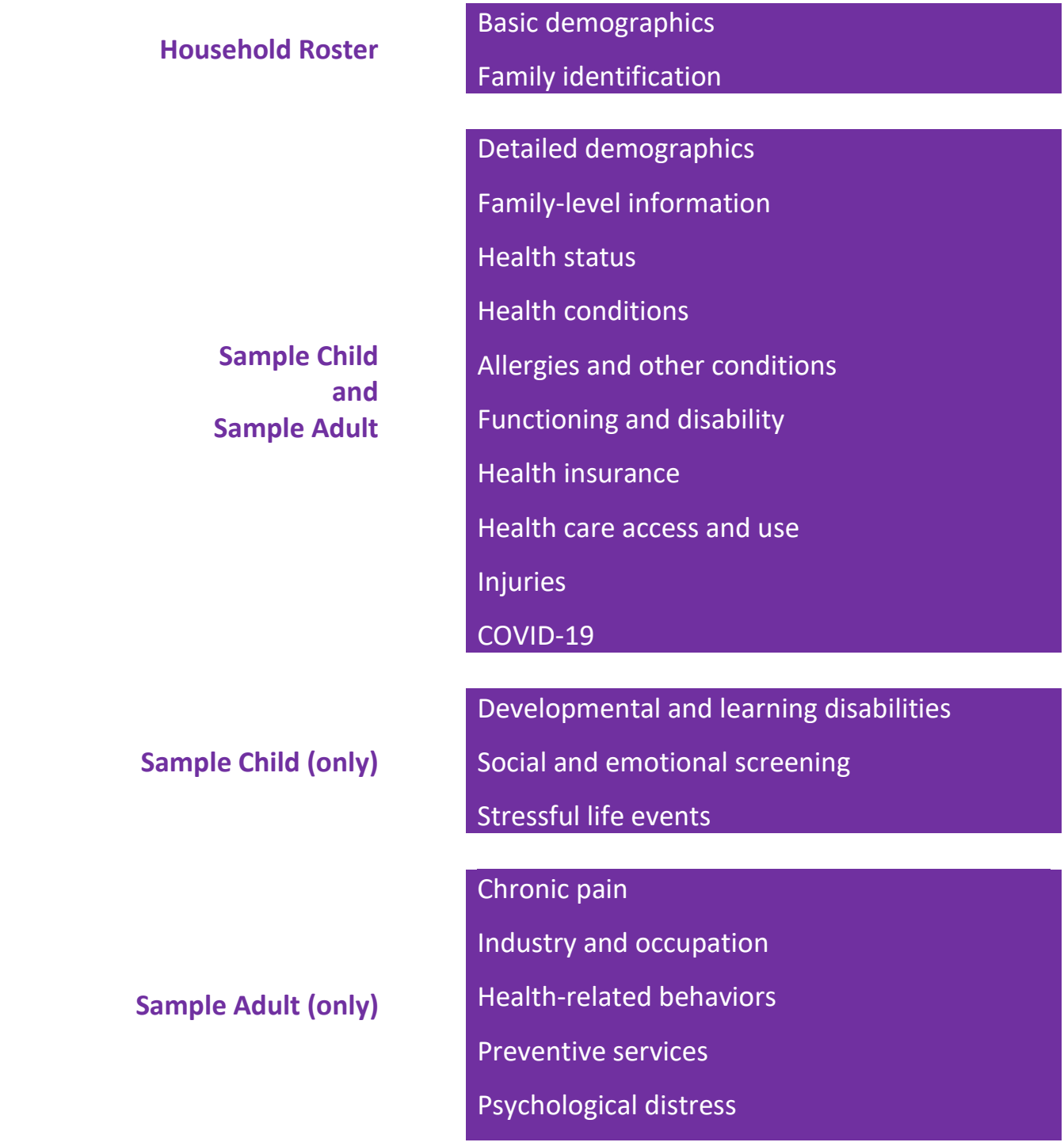
The redesigned NHIS questionnaire consists of three modules: (1) Household Roster; (2) Sample Adult Interview; and (3) Sample Child Interview.

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 interviewers can then proceed with the Sample Adult or Sample Child interview (if a child lives in the household).

Figure 3. Topic organization in the Household Roster, Sample Adult and Sample Child modules: NHIS 2021.



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 their 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 their 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. 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 2021 Questionnaire

A description of the 2021 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 2021 are:

I. Health Status and Conditions
self-reported health status, height, weight, pregnancy status, and the following health conditions: allergies, angina pectoris, anxiety disorder, arthritis, rheumatoid arthritis, gout, lupus, or fibromyalgia, asthma, cancer and cancer type, chronic fatigue syndrome, chronic obstructive pulmonary disease, cirrhosis or any other kind of long-term liver condition, cold or flu (longer than a month), coronary heart disease, dementia including Alzheimer's disease, depression, diabetes, dry mouth, epilepsy or seizure disorder, gestational diabetes, hepatitis (any type), hyperlipidemia, hypertension, immunosuppression, myocardial infarction, pre-diabetes, smell disorder, stroke, taste disorder, and weak or failing kidneys

II. Functioning and Disability
anxiety, cognition, communication, depression, hearing, mobility, self-care and upper body, social functioning (participation), vision, and age of disability onset

III. Chronic Pain
chronic pain and pain locations

IV. Health Care Access and Health Service Utilization
difficulty paying for health care; health care use (medical care visit, wellness visit, urgent care, emergency care, hospitalization, mental health care); hepatitis exposure; immunizations (COVID-19, flu, hepatitis A, hepatitis B, Pneumonia, tetanus, shingles); prescription medication; usual place for medical care; unmet medical, prescription, and mental health care due to cost; and work or volunteer activities in health care settings

V. Preventive Care
Aspirin use; clinical checks for arterial blood pressure, blood glucose, A1C level, and cholesterol; screenings for breast, cervical, colorectal, and prostate cancer

VI. Health-Related Behaviors
cigarette smoking, electronic cigarette use, and use of other tobacco products

VII. Mental Health
Kessler 6 nonspecific distress scale, and life satisfaction

VIII. Injuries
repetitive strain injuries, and sudden onset injuries

IX. COVID-19
COVID-19 testing, diagnosis and symptom severity, loss of taste and smell and recovery, delay in medical care, cancer care and treatment, caregiving received, social distancing at work, and social support

Sample Child health topics for 2021 are:

I. Health Status and Conditions
reported health status of the Sample Child, and the following health conditions: allergies, asthma, developmental conditions, diabetes, 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; health care use (medical care visit, wellness visit, urgent care, emergency care, hospitalization, mental health care); presence of a personal health care professional and receipt of care without parent/guardian; immunizations (COVID-19, flu); prescription medication; usual place for medical care; and unmet medical care and prescription use due to cost

IV. Behavioral and Mental health
Baby Pediatric Symptom Checklist (BPSC), bullying, social support, and stressful life events

V. Stressful Life Events
Experienced verbal abuse, had a parent who was incarcerated after birth, had unmet basic needs, lived with someone with a mental illness, or drug or alcohol problem, treated unfairly because of race or ethnic group, or sexual orientation, and witnessed or experienced neighborhood violence

VI. Injuries
Sudden onset injuries, and concussions

VII. COVID-19
COVID-19 testing, diagnosis and symptom severity, and delay in medical care

IV. Sponsors

Some 2021 NHIS content is sponsored by other federal agencies or other centers within the Centers for Disease Control and Prevention. 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 add content every year. Other sponsors add content for selected years.

NHIS Sustaining Sponsors

Cancer Control and Prevention

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 cancer treatment and care during the COVID-19 pandemic, and preventive screenings for colorectal, breast, cervical, and prostate cancers.

Immunizations and Employment in Health Care Settings

The National Center for Immunization and Respiratory Diseases at the Centers for Disease Control and Prevention (CDC/NCIRD) sponsored 24 Sample Adult questions about flu vaccination during pregnancy, COVID-19 vaccination (beginning in Q2), hepatitis, shingles, and two Sample Adult questions about working or volunteering in the health care industry. Beginning in Q3, five questions on COVID-19 vaccination were added for the Sample Child.

Noncigarette Tobacco Product 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.

Food Security and Food Stamp Benefits

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 sponsored a question about use of SNAP in the past 30 days. (Use of SNAP in the past year is part of the annual core content.)

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.

Taste and Smell

National Institute on Deafness and Other Communication Disorders at the National Institutes of Health (NIH/NIDCD) sponsored for 2021 four questions on COVID-related taste and smell, two questions on conditions related to taste and smell, and nine questions on difficulties with taste and smell for Sample Adults.

Other NHIS Sponsors

Age of Disability Onset

The Administration for Community Living at the Department of Health and Human Services (ACL/HHS) sponsored for 2020–2023 a Sample Adult question about the age of disability onset.

Chronic Fatigue Syndrome/ME

The National Center for Emerging and Zoonotic Infectious Diseases (CDC/NCEZID) sponsored for 2021 two Sample Adult questions on chronic fatigue syndrome (CFS) or myalgic encephalomyelitis (ME).

Diabetes

The National Center for Chronic Disease Prevention and Health Promotion at the Centers for Disease Control and Prevention (CDC/NCCDPHP) sponsored for 2021 seven Sample Adult questions about blood level check-up, unmet insulin needs, and stressed by diabetes.

Epilepsy

The National Center for Chronic Disease Prevention and Health Promotion at the Centers for Disease Control and Prevention (CDC/NCCDPHP) sponsored for 2021–2023 four Sample Adult questions on epilepsy.

Satisfaction With Life

The National Center for Chronic Disease Prevention and Health Promotion at the Centers for Disease Control and Prevention (CDC/NCCDPHP) and the Office of Disease Prevention at the National Institutes of Health (NIH/ODP) sponsored for 2021 two questions on satisfaction with life for Sample Adults and one question for Sample Children.

Social Distancing at Work and Work Arrangements

The National Center for Occupational Safety and Health at the Centers for Disease Control and Prevention (CDC/NIOSH) sponsored for 2020-20211 ten Sample Adult questions on social distancing measures at work and for 2021 nine Sample Adult questions on work arrangements, including scheduling and psychosocial job stressors.

Stressful Life Events

The National Center for Injury Prevention and Control at the Centers for Disease Control and Prevention (CDC/NCIPC) sponsored for 2021 two Sample Child questions on stressful life events.

V. Sample Sizes and Response Rates

When the NHIS sample was redesigned for 2016–2025, the base of approximately 58,800 addresses was expected to yield about 27,000 Sample Adult and 9,000 Sample Child interviews in roughly 35,000 households each year. Adjusting for response rate changes to the redesigned questionnaire implemented in 2019, the base sample is now expected to yield approximately 28,800 Sample Adult and 8,400 Sample Child interviews in 30,000 households annually. However, NHIS sample size may vary from year to year. Table 1 provides a breakdown of sample sizes for the various components of the 2021 NHIS public-use data release.

Table 1. Final sample sizes for the 2021 NHIS public-use data release

Interview Unit	2021 Sample
Households	30,673
Sample Adults	29,482
Sample Children	8,261

*** For 582 of the 29,482 Sample Adults on the 2021 sample, 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 below 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. Partials 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.

2021 Sample

Household Response Rate

For the 2021 Sample, the household response rate was calculated by dividing the number of interviewed households (n=30,673) by the sum of the number of interviewed households (n=30,673) and the number of nonresponding households (n=27,395). 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 the 2021 Sample was 52.8%.

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 2019-2021 NHIS should be presented separately.

Household Roster Completion

The completion of the Household Roster is defined as the enumeration of all 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 (n=33,325) by the number of eligible households (n=58,068). For the 2021 Sample, the Household Roster Completion Rate was 57.4%. Based on demographic information obtained from completed household rosters, there were 9,511 eligible Sample Children and 33,264 eligible Sample Adults in the 2021 Sample.

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 (n=8,261) by the number of eligible Sample Children from households with completed rosters (n=9,511). For the 2021 Sample, the Conditional Sample Child Response Rate was 86.9%.

The Final Sample Child Response Rate accounts for the Household Roster Completion Rate and is calculated by dividing the number of interviewed Sample Children (n=8,261) by the number of eligible Sample Children (n=9,511) from households with completed rosters, and then multiplying this quotient by the Household Roster Completion Rate (57.4%). In 2021, 1.4% of Sample Child interviews were sufficient partials.

For the 2021 Sample, the Final Sample Child Response Rate was 49.9%.

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 (n=29,482) by the number of eligible Sample Adults from households with completed rosters (n=33,264). For the 2021 Sample, the Conditional Sample Adult Response Rate was 88.6%.

The Final Sample Adult Response Rate is calculated by dividing the number of interviewed Sample Adults (n=29,482) by the number of eligible Sample Adults from households with completed rosters (n=33,264), and

then multiplying this quotient by the Household Roster Completion Rate (57.4%). In the 2021 Sample, 4.0% of Sample Adult interviews were sufficient partials.

For the 2021 Sample, the Final Sample Adult Response Rate was 50.9%.

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.

Summary of Household, Sample Adult, and Sample Child Response Rates Since 2019

The 2019-2021 counts for eligible and interviewed sample units used in the calculation of response rates for combined data years are shown in Table 2. Tables 3 and 4 present the conditional and unconditional response rates for the household, Sample Adult, and Sample Child modules for 2019-2021.

Table 2. Number of eligible and interviewed households, Sample Adults and Sample Children, National Health Interview Survey, 2019-2021

	Household			Sample Adult		Sample Child	
Year	Eligible	Roster Complete	Interviewed	Eligible	Interviewed	Eligible	Interviewed
2019	54,231	35,404	33,138	35,365	31,997	10,155	9,193
2020	43,280	23,694	21,930	23,694	21,153	6,626	5,790
2021	58,068	33,325	30,673	33,264	29,482	9,511	8,261

Table 3. Conditional response rates, National Health Interview Survey, 2019-2021

Year	Sample Adult	Sample Child
2019	90.5	90.5
2020	89.3	87.4
2021	88.6	86.9

Note: The Conditional Response Rate is calculated by dividing the number of interviewed by the number of eligible from households with completed rosters.

Table 4. Unconditional response rates, National Health Interview Survey, 2019-2021			
Year	**Household**	**Sample Adult**	**Sample Child**
2019	61.1	59.1	59.1
2020	50.7	48.9	47.8
2021	52.8	50.9	49.9
Note: This is the Final Response Rate, and it is calculated by dividing the number of interviewed by the number of eligible from households with completed rosters, and then multiplying this quotient by the Household Roster Completion.			

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. 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. In 2019 the updated approach for nonresponse adjustment used 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. In 2021, instead of using multilevel regression models to determine the nonresponse adjustment cells, Recursive Partitioning for Modeling Survey Data (RPMS) was used within R programming language (Toth, 2019). One characteristic of the RPMS is that this regression tree model can account for stratification and clustering as well as unequal probabilities of selection. A major advantage of RPMS is the automatic consideration and estimation of complex interactions.

Finally, the nonresponse adjusted weights are typically calibrated to U.S. Census Bureau population projections and American Community Survey (ACS) one-year estimates for age, sex, race and ethnicity, and since 2019, also for educational attainment, and Census division by Metropolitan Statistical Area (MSA) status. In 2020, housing tenure was added to the calibration step. For the 2021 survey year, the U.S. Census Bureau did not release

single-year ACS estimates by housing tenure, education level, and MSA by Division. Therefore, substitute calibration totals for these variables were obtained from the 2021 Current Population Survey (CPS) March Annual Social and Economic (ASEC) Supplement. These changes to the nonresponse adjustment approach and the calibration methods have the potential to impact comparisons of the weighted survey estimates over time.

2021 Weights

The Sample Adult and the Sample Child file each have a unique and separate final annual weights and variance estimation variables.

The Final Annual Weight should be used to generate national estimates. This weight includes the design, ratio, nonresponse and calibration adjustments.

The final annual weight in the Sample Adult file is WTFA_A.

The final annual weight in the Sample Child file is WTFA_C.

Further information on how to implement sampling weights in found in the section "Analyzing 2021 NHIS."

The paradata files do not contain weights. Depending on the analysis of the paradata, if weights are needed, they can be pulled from public-use data files.

Interim weights are those sampling weights that do not include the final standard calibration adjustment for age, sex, race and ethnicity, education, and Census division by MSA status raking to population control totals. Since 2020, interim weights are not included in the public use files and are available through the NCHS Research Data Center (RDC): https://www.cdc.gov/rdc/.

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 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 RDC:

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.

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

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 302 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 a type of cancer (e.g., LUNGAGE_A) and diabetes (DIBAGE_A) ("How old were you when you were diagnosed with [this condition]?") are top-coded to 85+ years.

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 annual restricted-use codebook available on the 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 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 is not 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). They can also be found in some sponsored sections of the questionnaire such as Food Security (FDS) and Food-Related Programs (FOO).

Replicate measures are used in analyses the same way as any other measure available in the Sample or Sample Child files. Analyses of replicate measures can be interpreted as an estimate of Sample Adults/Sample Children/persons who are in a family meeting a specific outcome or characteristic (e.g., percentage of persons aged 0–64 years who are in a family that is having problems paying medical bills).

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 2021 are available on the NHIS website:
https://www.cdc.gov/nchs/nhis/2021nhis.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.

Imputed income technical document (PDF): A description of the methodology for creating the 2021 NHIS imputed income variables.

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

Industry and Occupation (xlsx): List of the 4-digit U.S. Census Bureau Industry and Occupation codes corresponding to the 2020–2021 NHIS Employment and Occupation data.

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.

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

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 list of sponsoring agencies and associated sponsored questions in NHIS.

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 2021 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) List of questions fielded only during the second, third, and fourth quarters of 2021
c) 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
d) 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 panel 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 respondent indicated that the Sample Child is male, "she" if female or "they" if sex was not provided. 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 2021 NHIS data release includes files for the annual Sample Adult, annual Sample Child, Imputed Income (as a poverty ratio) for the Sample Adult and Sample Child, and Paradata.

The file names for the 2021 data release are listed in Table 5. Files corresponding to the 2021 NHIS have a two-digit suffix at the end that represents the survey year, e.g., 21 for 2021, or adult21. In years prior to 2019, separate files were available for household, family, and person level information due to the different survey design. In 2020, additional files were available for the Sample Adult (Partial and Longitudinal) due to the longitudinal component implemented that year but not continued for 2021. 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 (see, "Merging Survey Data and Imputed Income Files" in next section). Sample Adult and Sample Child files can also be merged with Paradata.

Sample Adult and Sample Child files

The 2021 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 2021 Imputed Income files for Sample Adult and Sample Child contain 10 imputations of family poverty ratio as both continuous and categorical 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 based on the first imputation were also added to the Sample Adult and Sample Child files for convenience of users who choose not to use multiply 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 from the other sets 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 collected 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 is described in the "Imputed Income Technical Document" available with the 2021 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 files, which contain 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.

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 2021 file releases on the NHIS website, under “Using the NHIS.”

File Names

The data 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 5 lists the names of data files, programs and documentation in the 2021 NHIS data release.

Data Documentation

Summary and Codebook PDFs provide accompanying documentation for the 2021 data 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 5. Data release files names for the Sample Adult, Sample Child, Imputed Income and Paradata files: NHIS 2021.

Type of file	Sample Adult Annual files	Sample Child Annual files	Imputed Income files	Paradata files
Data in column-delimited ASCII format	adult21.dat	child21.dat	Adultinc21.dat Childinc21.dat	paradata21.dat
Data in comma separated Values (CSV) format	adult21.csv	child21.csv	Adultinc21.csv Childinc21.csv	paradata21.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.pdf Childinc-summary.pdf	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 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 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. Examples of recodes based on annual core content include MARSTAT_A (current marital status of the Sample Adult) and SMKCIGST_A (cigarette smoking status). Additionally, a recode may be created when a continuous variable needs top- and/or bottom-coding for confidentiality reasons (for example, SCHDYMSSTC_C, school days missed due to illness or injury in the past 12 months). In some years, the NHIS may also contain content on the frequency of a health behavior in two parts: number and type of time unit. Recodes are then created to combine and standardize responses into a single time unit.

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.

F1*:* Indicates there is a 'help screen' available to the interviewers with definitions about the question topic that can be used for answering respondents' questions.

II. Analyzing 2021 NHIS

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).

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 examples below use the Sample Adult sampling weight (WTFA_A). For Sample Child analysis, use the Sample Child sampling weight (WTFA_C).

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 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.

Examples of Analysis and Weighting Procedures for Sample Adults

Tables 6 includes examples of types of research questions that analysts may use 2021 Sample Adult NHIS data for, and guidance on additional adjustments needed to the sampling weights when analyzing in combination with the 2019–2020 NHIS. Note that the 2020 NHIS included three different sampling weights for the Sample Adult, and these were designated for the analyses of the longitudinal sample, the partial sample (2020 sample only), and the combination of those two samples.

Analysts may apply some of these examples to Sample Child files and use instead the Sample Child annual weight (WTFA_C). For example, analysts may concatenate annual Sample Child files to increase sample size or to compare estimates over time (e.g., prepandemic versus during pandemic, or COVID-19 related measures since July 2020). Note that the Sample Child annual files did not include a longitudinal sample as done with Sample Adults in 2020, and any procedure described below referencing the longitudinal sample is not applicable to Sample Child files.

Table 6. Examples of Analysis and Weighting Procedures for Sample Adults—2021 NHIS

Analytic Goal	Example	Weight Variable and Modifications Needed	Variance Estimation
Produce official estimates for 2021	Percentage of adults ever told by doctor or other health professional that they had diabetes	Use WTFA_A	Use standard variance estimation variables PSTRAT and PPSU for 2021
Compare estimates between 2021 and 2020	Percentage of adults ever told by doctor or other health professional that they had diabetes, 2021 compared with 2020	Use WTFA_A for both 2021 and 2020	Use standard variance estimation variables PSTRAT and PPSU for 2021 and 2020
Produce estimates and frequency of content only available in quarters 2 through 4 of 2021	Percentage of adults who were ever vaccinated for COVID-19	Create a new weight variable where the observations in quarter 1 are assigned a value of zero, and for observations in quarters 2 to 4, multiply the weight (WTFA_A) by 4/3	Use standard variance estimation variables PSTRAT and PPSU for 2021
Pooling 2021 and 2020 to increase sample size [1]	Percentage of adults who use hearing aids	Concatenate the 2020 and 2021 files. Create a new weight using WTFA_A and divide it by 2	Use standard variance estimation variables PSTRAT and PPSU for 2021 and 2020

Table 6 Continued. Examples of Analysis and Weighting Procedures for Sample Adults—2021 NHIS

Analytic Goal	Example	Weight Variable and Modifications Needed	Variance Estimation
Produce estimates of COVID-19 content since July 2020. 18 months of data combined to increase sample size (7/2020 to 12/2021) [2]	Percentage of adults who delayed getting medical care due to the pandemic	Create a new weight variable where the observations in quarters 1 and 2 of 2020 are assigned a value of zero and observations in quarters 3 and 4 of 2020 are assigned the values of WTFA_A. All quarters of 2021 are used for the new weight variable from WTFA_A. Concatenate the 2020 and 2021 files and multiply the new weight by 2/3 in the combined file	Use standard variance estimation variables PSTRAT and PPSU for 2021 and 2020
Produce estimates of COVID-19 content from July 2020 through December 2021 as 6-month intervals. Compare semi-annual intervals (trends).[3]	Percentage of adults who delayed getting medical care due to the pandemic: July-December 2020, January-June 2021, and July-December 2021	Create a new weight variable where the observations in quarters 1 and 2 of 2020 are assigned a value of zero; and for each remaining 6-month interval, multiply the weight by 2.	Use standard variance estimation variables PSTRAT and PPSU for 2021 and 2020
Pooling 2019, 2020 and 2021 data to increase sample size.[4]	Percentage of adults ever told by doctor or other health professional that they had a stroke	Use WTFA_A in 2019 and 2021, and the partial weight WTSA_P for 2020. Create a new weight using these weight variables and divide it by 3	Use standard variance estimation variables PSTRAT and PPSU for 2019-2021
Examine estimates between 2019 and quarters 2 through 4 of 2021 for a subgroup of adults.[5]	Percentage of adults aged 40 and older taking low-dose aspirin daily, pre-pandemic versus during the pandemic	Use 2019 weight (WTFA_A) as is, and for 2021, multiply the weight (WTFA_A) by 4/3	Use standard variance estimation variables PSTRAT and PPSU for 2019 and 2021, Subset adults aged 40+ using the AGEP_A variable

[1] See scenario 1 under "Appending and Merging Sample Adult files" for example using SAS or STATA.
[2] See scenario 2 under "Appending and Merging Sample Adult files" for example using SAS or STATA.
[3] See scenario 3 under "Appending and Merging Sample Adult files" for example using SAS or STATA.
[4] See scenario 4 under "Appending and Merging Sample Adult files" for example using SAS or STATA.
[5] See scenario 5 under "Appending and Merging Sample Adult files" for example using SAS or STATA.

Appending and Merging Sample Adult files

This section provides sample code in SAS and STATA for appending and merging Sample Adult data files for conducting data analysis. The sample code follows examples listed in Table 6.

Scenario 1: Concatenating (pooling) 2020 and 2021 Sample Adult data to increase sample size

This example illustrates how to create a dataset that pools data from 2020 and 2021 to increase sample size and produce a cross-sectional estimate (e.g., use hearing aids). An analyst can pool 2020 and 2021 data when interested in increasing the precision of an estimate among a population that might have a small sample size with a single year of data. For this example, a new file called ADULT20_21 will include all Sample Adults in 2020 and 2021 data. It will be derived from the data files, adult20 and adult21. This new file will only contain the variables of interest: use hearing aid, variance structure variables, and sampling weights from the adult20 and adult21 files. The new weight WTFA_ADJ will be created to account for 2 years of data (i.e., that is dividing each sample weight by the number of years that are being pooled).

Example using SAS

```
*Create temporary datasets of 2020 and 2021 data by selecting weight, analytic variables and variance
structures (PSTRAT and PPSU);
DATA TEMPADULT20;
SET ADULT20; *2020 Sample Adult file;
KEEP WTFA_A PSTRAT PPSU HEARAID_A;
RUN;

DATA TEMPADULT21;
SET ADULT21; *2021 Sample Adult file;
KEEP WTFA_A PSTRAT PPSU HEARAID_A;
RUN;

*Concatenate the temporary 2020 and 2021 datasets;
DATA ADULT20_21;
SET TEMPADULT20 TEMPADULT21;
WTFA_ADJ=WTFA_A/2; *Divide the weight by 2;
RUN;
```

Example using Stata

```
*Create temporary datasets of 2020 and 2021 data by selecting weight, analytic variables and variance
structures (PSTRAT and PPSU)
*Adult20 – 2020 Sample Adult file
use adult20
keep wtfa_a pstrat ppsu hearaid_a
save tempadult20
```

```
*Adult21 – 2021 Sample Adult file
use adult21
keep wtfa_a pstrat ppsu hearaid_a
save tempadult21

*Concatenate temporary 2020 and 2021 datasets
append using tempadult20
gen wtfa_adj=.
*Divide the weight by 2;
replace wtfa_adj=wtfa_a/2
save adult20_21
```

Scenario 2: 18 months of data combined. Produce estimates of COVID-19 content since July 2020 (7/2020 to 12/2021) to increase sample size

This example illustrates how to create a dataset that pools data from the third and fourth quarters of 2020 and a full year of data from 2021 to increase sample size and produce cross-sectional estimates (e.g., delayed getting medical care due to the pandemic). An analyst can pool 2020 and 2021 data when interested in increasing the precision of an estimate. For this example, a new file called NEWADULT20_21 will include all Sample Adults in 2020 and 2021, and it will be derived from the data files, adult20 and adult21.

A new weight variable will be created in a temporary 2020 file from WTFA_A called NEWWT that assigns Sample Adults for the first two quarters in 2020 a weight of zero and Sample Adults from the last two quarters in 2020 the values from WTFA_A. A new weight variable will be created in a temporary 2021 file that renames WTFA_A to NEWWT. A final file, NEWADULT20_21, is created by concatenating the temporary 2020 and 2021 data files. The final file will only contain the variables of interest: delaying care due to COVID-19, variance structure variables, and the new sampling weight for the 18 months analysis. The new weight, NEWWT_ADJ, is created to account for 18 months of data and it is derived from multiplying NEWWT by 2/3.

Example using SAS

```
*Create temporary datasets of 2020 data by selecting weight, analytic variables and variance structures
(PSTRAT and PPSU);
DATA NEWADULT20;
SET ADULT20; *2020 Sample Adult file;
KEEP PSTRAT PPSU DLYCARE_A NEWWT;
IF INTV_QRT IN (1, 2) THEN NEWWT = 0; *For the first and second quarters in 2020, assign weights of
zero;
ELSE NEWWT=WTFA_A;
RUN;

*Create a temporary subset of 2021 data by selecting weight, analytic variables, and variance structures
(PSTRAT and PPSU);
DATA NEWADULT21;
SET ADULT21; *2021 Sample Adult file;
KEEP PSTRAT PPSU DLYCARE_A NEWWT;
```

```
NEWWT=WTFA_A; *Rename WTFA_A to NEWWT to match the variable name for sampling weight;
RUN;

*Concatenate temporary 2020 and 2021 datasets;
DATA NEWADULT20_21;
SET NEWADULT20 NEWADULT21;
KEEP PSTRAT PPSU DLYCARE_A  NEWWT_ADJ;
NEWWT_ADJ=NEWWT * 2/3; *For 18 months of data, multiply the weight by 2/3;
RUN;
```

Example using Stata

```
* Create temporary datasets of 2020 data by selecting weight, analytic variables and variance structures
(PSTRAT and PPSU)
use adult20
keep pstrat ppsu dlycare_a intv_qrt wtfa_a
gen newwt=.
*For the first and second quarters in 2020, assign weights of zero
replace newwt=0 if intv_qrt == 1 | intv_qrt == 2
replace newwt=wtfa_a if intv_qrt == 3 | intv_qrt == 4
drop wtfa_a intv_qrt
save newadult20

*Adult21 – 2021 Sample Adult file
*Create a temporary subset of 2021 data by selecting weight, analytic variables, and variance structures
(PSTRAT and PPSU)
use adult21
keep pstrat ppsu dlycare_a wtfa_a
*Rename wtfa_a to newwt to match the variable name for sampling weight
gen newwt=.
replace newwt=wtfa_a
drop wtfa_a
save newadult21

*Concatenate temporary 2020 and 2021 datasets
append using newadult20
keep pstrat ppsu dlycare_a  newwt
gen newwt_adj=.
*For 18 months of data, multiply the weight by 2/3
replace newwt_adj=newwt * 2/3
drop newwt
save newadult20_21
```

Scenario 3: Comparing semi-annual intervals (trends): Produce estimates of COVID-19 content from July 2020 through December 2021 as 6-month intervals

This example illustrates how to create one dataset to examine semi-annual intervals (trends) over an 18-month period of data collection. Data from the third and fourth quarters of 2020 and a full year of data from 2021 will be used in this example to produce cross-sectional 6-month interval estimates (e.g., delayed getting medical care due to the pandemic).

For this example, one new file will be created called NEWADULT20_21. This new file will only contain the variables of interest: delaying care due to COVID-19, variance structure variables, modified sampling weights from the adult20 and adult21 files and an indicator for each 6-month interval (TIME6M). For each 6-month interval, a NEWWT will be created by multiplying the weight (WTFA_A) by 2 to account for only using 6-months of data. Estimates in this dataset should only be examined by 6-month intervals, which is indicated by TIME6M variable (i.e., value 1 is the first 6-month period, value 2 is the second 6-month period, and value 3 is the third 6-month period). The annual weight has been adjusted to treat 6 months of data so that weighted frequencies for a particular measure (e.g., delaying care due to COVID-19) will reflect the number of people who have delayed care due to COVID-19 in the U.S. civilian noninstitutionalized population. Using WTFA_A without adjustment for a 6-month analysis will result in population estimates that are too low and will not reflect the number of people in the U.S. civilian noninstitutionalized population that have delayed care due to COVID-19.

Example using SAS

```
*Create temporary datasets of 2020 data by selecting weight, analytic variables and variance structures (PSTRAT and PPSU);
DATA NEWADULT20;
SET ADULT20; *2020 Sample Adult file;
KEEP PSTRAT PPSU DLYCARE_A NEWWT TIME6M;
IF INTV_QRT IN (1, 2) THEN NEWWT = 0; *For the first and second quarters in 2020, assign weights of zero;
ELSE NEWWT=WTFA_A * 2; *For 6-months of data, multiply the weight by 2;
IF INTV_QRT IN (3, 4) THEN TIME6M=1; *This value indicates the first 6-month time period that data was collected;
RUN;

*Create temporary dataset of 2021 data by selecting weight, analytic variables and variance structures (PSTRAT and PPSU);
DATA NEWADULT21;
SET ADULT21; *2021 Sample Adult file;
KEEP PSTRAT PPSU DLYCARE_A NEWWT TIME6M;
IF INTV_QRT IN (1, 2) THEN TIME6M=2; *For the first and second quarters of 2021, this is the second 6-month time period that data was collected;
IF INTV_QRT IN (3, 4) THEN TIME6M=3; *For the third and fourth quarters of 2021, this is the third 6-month time period that data was collected;
NEWWT = WTFA_A *2; *For 6-months of data, multiply the weight by 2;
RUN;

*Concatenate temporary 2020 and 2021 datasets;
DATA NEWADULT20_21;
```

```
SET NEWADULT20 NEWADULT21;
KEEP PSTRAT PPSU DLYCARE_A NEWWT TIME6M;
RUN;
```

Example using Stata

```
*Create temporary datasets of 2020 data by selecting weight, analytic variables and variance structures (PSTRAT and PPSU)
use adult20
keep pstrat ppsu dlycare_a  intv_qrt wtfa_a
gen newwt=.
*For the first and second quarters in 2020, assign weights of zero
replace newwt=0 if intv_qrt == 1 | intv_qrt == 2
*For 6-months of data, multiply the weight by 2
replace newwt=wtfa_a * 2 if intv_qrt == 3 | intv_qrt == 4
*This value indicates the first 6-month time period that data was collected
gen time6m=1 if intv_qrt == 3 | intv_qrt == 4
drop wtfa_a intv_qrt
save newadult20

*Create temporary dataset of quarters 1 and 2 of 2021 data by selecting weight, analytic variables and variance structures (PSTRAT and PPSU)
use adult21
keep pstrat ppsu dlycare_a intv_qrt wtfa_a
gen time6m=.
*For the first and second quarters of 2021, this is the second 6-month time period that data was collected
replace time6m=2 if intv_qrt == 1 | intv_qrt == 2
*For the third and fourth quarters of 2021, this is the third 6-month time period that data was collected
replace time6m=3 if intv_qrt == 3 | intv_qrt == 4
*For 6-months of data, multiply the weight by 2
gen newwt = wtfa_a * 2
drop wtfa_a intv_qrt
save newadult21

*Concatenate temporary 2020 and 2021 datasets
append using newadult20
keep pstrat ppsu dlycare_a newwt time6m
save newadult20_21
```

Scenario 4: Concatenating (pooling) 3 years of data: 2019, 2020, and 2021 to increase sample size (and excluding 2020 observations from Sample Adults interviewed both in 2019 and 2020)

This example illustrates how to create a dataset that pools data from 2019, 2020, and 2021 to increase sample size and produce a cross-sectional estimate (e.g., ever told by a doctor or other health professional that they had a stroke). An analyst can pool 2019, 2020 and 2021 data when interested in increasing precision of an estimate among a population that might have a small sample size with a single year of data, especially when examining among subgroups (e.g., adults younger than 65 years of age, living in nonmetropolitan areas).

This analysis excludes the responses from the second interview of the longitudinal sample. For this example, a new file called NEWADULT19_20_21 will include all Sample Adults in 2019, Sample Adults only interview in 2020 (and not part of the longitudinal sample), and all Sample Adults in 2021. It will be derived from the data files, adult19, adult20, adultpart20 and adult21. A temporary dataset will be created from the adult19 file that will only contain the 2019 variables for stroke, variance structure and the sampling weight. The adultpart20 file will be merged with the adult20 file to create a temporary 2020 dataset that will only contain the variables of interest: stroke and the variance structure variables from the adult20 file, and the sampling weight from the adultpart20 file. The adultpart20 file contains the weight variable for Sample Adults in 2020 sample who were not part of the longitudinal sample. A third temporary dataset will be created from the adult21 file that will only contain the 2021 variables for stroke, variance structure and the sampling weight. The 2019, 2020, and 2021 temporary files will be combined into a new dataset, NEWADULT19_20_21. It will have a new weight (NEWWT_ADJ) where the annual weight (WTFA_A) will apply to 2019 and 2021 cases, and the partial weight (WTSA_P) will apply to 2020 cases. This weight will also be adjusted to account for 3 years of data (i.e., that is dividing each sample weight by the number of years that are being pooled).

Example using SAS

```
*Create a temporary subset of 2019 data by selecting weight, analytic variables, and variance structures (PSTRAT and PPSU);
DATA NEWADULT19;
SET ADULT19; *2019 Sample Adult file;
KEEP WTFA_A PSTRAT PPSU STREV_A NEWWT;
NEWWT=WTFA_A;
RUN;

*Create temporary subsets of 2020 data by selecting household ID (HHX), weight, analytic variables and variance structures (PSTRAT and PPSU);
DATA TEMPADULT20;
SET ADULT20; *2020 Sample Adult file;
KEEP HHX PSTRAT PPSU STREV_A NHHX;
NHHX=HHX; *Rename HHX to NHHX to match with household ID in 2020 Sample Adult partial file;
RUN;

DATA TEMPPART20;
SET ADULTPART20; *2020 Sample Adult partial file;
KEEP HHX_2020 WTSA_P NHHX;
NHHX=HHX_2020; *Rename HHX_2020 to NHHX to match with household ID in 2020 Sample Adult file;
RUN;
```

```
*Sort each temporary dataset by the merge variable;
PROC SORT DATA=TEMPADULT20;
BY NHHX;
PROC SORT DATA=TEMPPART20;
BY NHHX;
RUN;

*Merge the two temporary 2020 Sample Adult files;
DATA NEWADULT20;
MERGE TEMPADULT20 TEMPPART20;
BY NHHX;
IF WTSA_P > 0; *Keep records with a value in the partial weight;
NEWWT=WTSA_P; *Rename WTSA_P to NEWWT to match the variable name for sampling weight when
combining files;
RUN;

*Create a temporary subset of 2021 data by selecting weight, analytic variables, and variance structures
(PSTRAT and PPSU);
DATA NEWADULT21;
SET ADULT21; *2021 Sample Adult file;
KEEP WTFA_A PSTRAT PPSU STREV_A NEWWT;
NEWWT=WTFA_A;
RUN;

*Concatenate temporary 2019, 2020, and 2021 datasets;
DATA NEWADULT19_20_21;
SET NEWADULT19 NEWADULT20 NEWADULT21;
KEEP PSTRAT PPSU STREV_A NEWWT_ADJ;
NEWWT_ADJ=NEWWT/3; *Divide the new weight by 3 for the three years of data being combined.
Otherwise, weighted estimates of totals will be higher than the estimated total U.S. civilian
noninstitutionalized population;
RUN;
```

Example using Stata

```
*Create a temporary subset of 2019 data by selecting weight, analytic variables, and variance structures
(PSTRAT and PPSU)
use adult19
keep wtfa_a pstrat ppsu strev_a
*Rename wtfa_a to newwt to match the variable name for sampling weight when combining files
gen newwt=.
replace newwt=wtfa_a
drop wtfa_a
save newadult19

*Create temporary subsets of 2020 data by selecting household ID (HHX), weight, analytic variables and
variance structures (PSTRAT and PPSU)
*Adult20 – 2020 Sample Adult file
```

```
use adult20
keep hhx pstrat ppsu strev_a
*Rename hhx to nhhx to match with household ID in 2020 Sample Adult partial file;
gen nhhx=""
replace nhhx=hhx
save tempadult20

*Adultpart20 – 2020 Sample Adult partial file
use adultpart20
*Rename hhx_2020 to nhhx to match with household ID in 2020 Sample Adult file
gen nhhx=""
replace nhhx=hhx_2020
keep hhx_2020 wtsa_p nhhx
save temppart20

*Sort each temporary dataset by the merge variable
use tempadult20
sort nhhx
save tempadult20, replace

use temppart20
sort nhhx
save temppart20, replace

*Merge the two temporary 2020 Sample Adult files
use tempadult20
merge 1:1 nhhx using temppart20
*Keep records with a value in the partial weight
keep if wtsa_p !=.
*Rename wtsa_p to newwt to match the variable name for sampling weight when combining files
gen newwt=.
replace newwt=wtsa_p
drop nhhx
save newadult20

*Concatenate 2019 and 2020 datasets
append using newadult19
keep pstrat ppsu strev_a newwt
save newadult19_20

*Create a temporary subset of 2021 data by selecting weight, analytic variables, and variance structures
(PSTRAT and PPSU)
use adult21
keep wtfa_a pstrat ppsu strev_a
*Rename wtfa_a to newt to match the variable name for sampling weight when combining files
gen newwt=.
replace newwt=wtfa_a
drop wtfa_a
save newadult21
```

```
*Concatenate 2019, 2020 and 2021 datasets
append using newadult19_20
keep pstrat ppsu strev_a newwt
*Divide the new weight by 3 for the three years of data being combined. Otherwise, weighted estimates
of totals will be higher than the estimated total U.S. civilian noninstitutionalized population
gen newwt_adj=.
replace newwt_adj=newwt/3
drop newwt
save newadult19_20_21
```

Scenario 5: Examine estimates between 2019 and quarters 2 through 4 of 2021 for a subgroup of adults

This example illustrates how to create one dataset to examine differences between 2019 and quarters 2 through 4 of 2021. Data from a full year of 2019 and three quarters of 2021 are used in this example to compare the estimates (e.g., use of low-dose aspirin) before the COVID-19 pandemic versus during the COVID-19 pandemic.

For this example, one new file will be created called NEWADULT19_21. For the 2019 data, the weight is renamed to NEWWT to match the weighting variable used for the 2021 estimates. For 2021 data, the first quarter will be assigned weights of zero and a NEWWT will be created by multiplying the weight (WTFA_A) by 4/3 to account for only using 3 quarters (9 months) of data. A variable to indicate which time period the data was collected is also included in the file (TIMEPER). Estimates in this dataset should only be examined by the time period variable. This example also keeps the variable for age (AGEP_A) since the question about low-dose aspirin use was only collected among adults aged 40 and older, and this variable is needed to subset the analysis to that subsample of adults (For information about subsetting analyses, see section “Variance Estimation for Subsetted Data Analysis”)

Example using SAS

```
*Create a temporary subset of 2019 data by selecting weight, analytic variables, and variance structures
(PSTRAT and PPSU);
DATA NEWADULT19;
SET ADULT19; *2019 Sample Adult file;
KEEP PSTRAT PPSU AGEP_A ASPMEDEV_A NEWWT TIMEPER;
NEWWT=WTFA_A; *Rename WTFA_A to NEWWT to match the variable name for sampling weight;
TIMEPER=1; *Data from the first time period (2019);
RUN;

*Create a temporary subset of 2021 data by selecting weight, analytic variables, and variance structures
(PSTRAT and PPSU);
DATA NEWADULT21;
SET ADULT21; *2021 Sample Adult file;
KEEP PSTRAT PPSU AGEP_A ASPMEDEV_A NEWWT TIMEPER;
IF INTV_QRT = 1 THEN NEWWT = 0; *For the first quarter in 2021, assign weights of zero;
ELSE NEWWT=WTFA_A * 4/3; *For 9-months of data, multiply the weight by 4/3;
IF INTV_QRT in (2, 3, 4) then TIMEPER=2; *Data from the second time period (2021);
RUN;
```

```
*Concatenate temporary 2019 and 2021 datasets;
DATA NEWADULT19_21;
SET NEWADULT19 NEWADULT21;
KEEP PSTRAT PPSU AGEP_A ASPMEDEV_A NEWWT TIMEPER;
RUN;
```

Example using Stata

```
*Create a temporary subset of 2019 data by selecting weight, analytic variables, and variance structures
(PSTRAT and PPSU)
use adult19
keep pstrat ppsu aspmedev_a wtfa_a agep_a
gen newwt=.
*Rename wtfa_a to newwt to match the variable name for sampling weight
replace newwt=wtfa_a
*Data from the first time period
gen timeper=1
drop wtfa_a
save newadult19

*Create a temporary subset of 2021 data by selecting weight, analytic variables, and variance structures
(PSTRAT and PPSU)
use adult21
keep pstrat ppsu agep_a aspmedev_a wtfa_a intv_qrt
gen newwt=.
*For the first quarter in 2021, assign weights of zero
replace newwt = 0 if intv_qrt == 1
*For 9-months of data, multiply the weight by 4/3
replace newwt = wtfa_a * 4/3 if intv_qrt == 2 | intv_qrt == 3 | intv_qrt == 4
*Data from the second time period
gen timeper=2
drop wtfa_a intv_qrt
save newadult21

*Concatenate temporary 2019 and 2021 datasets
append using newadult19
keep pstrat ppsu agep_a aspmedev_a newwt timeper
save newadult19_21
```

Merging Survey Data and Paradata Files

Data users can merge the Paradata file with the Sample Adult file (or the Sample Child file) to explore associations between a wide range of methodological measures and survey data. To merge 2021 files, use variable HHX as the unique identifier between the two files. Sample code in SAS and STATA is provided below to illustrate merging the Sample adult file with the Paradata file.

Note that data files should be merged within the same year before combining (pooling) data files for multiple years.

Example using SAS

```
PROC SORT DATA=ADULT21;
BY HHX;
RUN;
PROC SORT DATA=PARADATA21;
BY HHX;
RUN;

/* creates a new file with Sample Adult and Paradata variables for each household */
DATA ADULT21_PLUS_PARA;
MERGE ADULT21 PARADATA21;
BY HHX;
RUN;
```

Example using STATA

```
*Change working directory to location of data files
cd c:\nhis2021\

use adult21
sort hhx
save adult21, replace

use paradata21
sort hhx
save paradata21, replace

merge 1:1 hhx using adult21
save adultpara21
* Creates a new file with Sample Adult and Paradata variables for each household
```

Merging Survey Data and Imputed Income Files

Data users can merge the Sample Adult file (or the Sample Child file) with their respective Imputed Income file to apply the imputation variable to the analyses for the appropriate calculation of standard error of the imputed variable. Variable HHX is the unique identifier between the two files.

The following code (in SAS, SUDAAN, and STATA) is for merging the Sample Adult data file and the Sample Adult imputed income file and conducting an analytic procedure (i.e., logistic regression using survey data). The variables used in this example were recoded as illustrated in Table 7. The analytic example, using the 2021 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 7. Variables in the imputed income example.

Survey question	Original variable name	Original values	Recoded variable name	Recoded values
Imputation number	IMPNUM_A	Values 01-10	(not recoded)	(not recoded)
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. 0.00–0.49 2. 0.50–0.74 3. 0.75–0.99 4. 1.00–1.24 5. 1.25–1.49 6. 1.50–1.74 7. 1.75–1.99 8. 2.00–2.49 9. 2.50–2.99 10. 3.00–3.49 11. 3.50–3.99 12. 4.00–4.49 13. 4.50–4.99 14. 5.00 or greater	(not recoded)	(not recoded)

Example 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 LOGITIC, 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_A 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 2021 NHIS Data Release page. This example uses the libname 'NHIS.'

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

DATA IMPINC;
SET NHIS.ADULTINC21;
RENAME IMPNUM_A= _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.ADULT21 OUT=ADULT21;
BY HHX;
RUN;

DATA NHIS21;
MERGE ADULT21 (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= NHIS21;
BY _IMPUTATION_;
RUN;
/*The survey analytic procedure (PROC SURVEYLOGISTIC) is used to account for the complex sampling
design of NHIS. */
/*WELLCHK is a recode from WELLNESS_A and HEALTH is a recode from PHSTAT_A*/
PROC SURVEYLOGISTIC DATA=NHIS21;
```

```
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 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.ADULTINC21 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 2021 NHIS Data Release page. This example uses the libname 'NHIS.'

```
PROC SORT DATA=nhis.ADULT21; *Sorting by HHX;
BY HHX;
RUN;

/* The following macro creates 10 separate imputation datasets from the NHIS.ADULTINC21 multiply
imputed file. It then sorts them by the merge variable HHX and merges each with the NHIS.ADULT21
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.ADULTINC21;
```

```
WHERE IMPNUM_A= &I;
RUN;

PROC SORT DATA= IMPINC&I; *Sorting the 10 imputed datasets by HHX;
BY HHX;
RUN;
DATA NHIS21_&I; *Creating 10 separate analytic datasets;
MERGE NHIS.ADULT21 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= NHIS21_&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. Variable HEALTH in this example is coded 0/1 */

PROC RLOGIST  DATA = NHIS21_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 waldffmt=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 using STATA

Import data files into Stata. See the sample Stata *.do* statements provided on the 2021 NHIS Data Release page.

```
*Change working directory to location of data files
cd c:\nhis2021\

use adult21
sort hhx
save nhis21

use adultinc21
sort hhx
save impinc

use nhis21
merge 1:m hhx using "impinc"
append using "nhis21"

* Rename the NHIS imputation number identifier impnum_a to _mi_m

replace impnum_a=0 if impnum_a==.
save nhis21_mi, replace

*set data to mi svyset

mi import flong, m(impnum_a) id(hhx)

*set dataset to multiple imputation for survey

mi svyset [pweight=wtfa_a], strat(pstrat) psu(ppsu) singleunit(centered)
save nhis21_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. The dependent variable, health, is coded as 0 or 1.

mi estimate: svy: logistic health ratcat_a wellchk
*odds ratios
mi estimate, or: svy: logistic health ratcat_a wellchk
```

Appending Sample Adult and Sample Child Files

The 2021 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 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 8.

Table 8. 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 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 PERSON21;
SET ADULT21 CHILD21;
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 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 child21
keep NOTCOV_C HHSTAT_C ppsu pstrat WTFA_C AGEP_C
save childvars

use adult21
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
```

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.

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.

Below are examples for subsetting NHIS data using SUDAAN, Stata, SPSS, SAS, and R software packages for illustrative purposes. 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 or African American only and SEX_A=2 is female.

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 Black or African American women;”
```

Using the full dataset with the SUBPOPN statement in this example would constrain this analysis to Black or African American women only (RACEALLP_A = 2 for Black or African American 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).

III. Questionnaire and Codebook Section Acronyms

Table 9. Acronym definition of sections in the questionnaire and codebooks: 2021 NHIS

Section	Section Description
ACC	Access to Care
ADO	Age of Disability Onset
ALG	Allergies
ANX	Anxiety
AST	Asthma
BCK	Back (Paradata)
BEH	Behavior
BLY	Bullying
BMI	Current pregnant, height, weight
BSC	Baby Pediatric Symptom Checklist
CAN	Cancer
CGR	Caregiving Received
CHL	Cholesterol
CIG	Cigarettes and E-cigarettes
CNV	Cancer COVID-19
COG	Cognition
COM	Communication
CON	Other Chronic Conditions
CVC	Cardiovascular Conditions
CVD	Positive COVID-19 Diagnosis
DEP	Depression
DIB	Diabetes
DLD	Developmental and Learning Disabilities
EMD	Detailed Employment
EMP	Employment
EPI	Epilepsy
FDS	Food Security
FEM	Employment of Family Members
FLG	Flags
FOO	Food Related Programs
FRT	Front
GEN	General
HEA	Hearing
HHC	Household Composition
HIS	Health Status
HOU	Housing
HYP	Hypertension
IDN	Identifier
IMM	Immunization (Sample Child)
IMS	Immunization (Sample Adult)
INC	Family Income
INJ	Injury
INS	Health Insurance
ISN	Immunosuppression

Table 9 Continued. Acronym definition of sections in the questionnaire and codebook: 2021 NHIS	
Section	**Section Description**
JOB	Work Arrangements
LNK	Linkage
LS1	Satisfaction with Life (Sample Child)
LS1/LS2/LS3	Satisfaction with Life (Sample Adult)
MAR	Marital Status
MHC	Mental Health Care
MOB	Mobility
NAT	Nativity
ORN	Sexual Orientation
OTB	Other Tobacco
PAI	Chronic Pain
PAR	Parent Demographics
PAY	Difficulty Paying for Health Care
PMD	Prescription Medication
PRV	Preventive Screening
RCN	Rotating Conditions
REL	Relationship of Children to Parents
REP	Repetitive Strain Injury
SCH	Schooling
SDW	Social Distancing at Work
SLE	Stressful Life Events
SOS	Social Support
SPD	Serious Psychological Distress
TBI	Concussion - Lifetime
TEL	Telephone Use
TSC	Taste and Smell - COVID
TSM	Taste and Smell
UPP	Upper Body, Motor Skills and Self-Care
UTZ	Utilization
VET	Veterans Status
VFY	Verification
VIS	Vision

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 10. Sample Adults were also asked about the intake of medication to treat diabetes, hypertension and high cholesterol, type of diabetes, 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 10. Annual core content measures of health conditions about the Sample Adult, by questionnaire section and reference periods.

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

Rotating Content

Rotating every three years starting in 2021 is content about allergies and other conditions.

The ALG section of the Sample Adult module included content about respiratory, food, and skin allergies. Sample Adults were asked whether they experience hay fever, seasonal or year-round allergies, an allergy to one or more foods, and an itchy rash due to eczema or atopic dermatitis. Those experiencing allergy symptoms were asked whether a doctor or other health care professional had ever told them that they have these respective allergies. Similar content is also available for the Sample Child.

The RCN section included questions about the diagnosis of additional chronic conditions. These questions asked whether a doctor or other health care professional had ever told the Sample Adult that they had:

- weak or failing kidneys
- hepatitis (any type)
- cirrhosis or any other kind of long-term liver condition

The questions began fielding in July 2020 as emerging content in response to the coronavirus disease 2019 (COVID-19) pandemic. Adults with kidney and liver disease are associated with experiencing severe COVID-19 (CDC, 2021).

Emerging Content

In response to the COVID-19 pandemic, two questions that asked Sample Adults about their immunosuppression status has been fielding since July of 2020 in the ISN section of the Sample Adult module as emerging content. Adults with an immunocompromised state are more likely to get severely ill from COVID-19 (CDC, 2021). Sample Adults were asked whether a doctor or other health professional had told them that:

- their prescription medication or any medical treatments in the past 12 months would weaken the immune system
- they currently have a condition that weakens the immune system

Sponsored Content

NIDCD sponsored nine questions in the TSM section of the Sample Adult module about smell and taste. Regarding their olfactory sense, Sample Adults were asked if in the past 12 months they had difficulty with their sense of smell or ability to detect odors, if they sometimes smell an unpleasant, bad, metallic, or burning odor when nothing was there, and to rate their current ability to smell now as compared to when they were younger (i.e., five years ago for those younger than 30 years, and at age 25 for those aged 30 and over). Regarding their sense of taste, Sample Adults were asked if in the past 12 months they have had difficulty with their ability to taste sweet, sour, salty, or bitter foods and drinks, and if they have had an unwanted taste or other sensation in their mouth that does not go away. They were also asked to rate their ability to taste sweet, sour, salty, or bitter foods as well as flavors (such as chocolate, vanilla, or strawberry) now as compared to when they were younger. Lastly, Sample Adults who had at least moderate difficulty with their sense of smell or taste, their ability to smell or taste was worse than when they were younger, or sometimes experienced phantosmia or dysgeusia were

asked if they had ever talked to a doctor or other health professional about their problem with their ability to taste or smell, and if so, the last time they had done so.
NIDCD also sponsored two questions in the RCN section that asked Sample Adults whether they had a cold or flu for longer than a month in the past 12 months, and if they experienced persistent dry mouth in the past 12 months. These questions are among several measures included in the 2021 NHIS about smell and taste disorders in adults. See also the description of the TSC section under COVID-19.

NCCDPHP and NIDDK sponsored eight questions 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 how long after diagnosis they began taking insulin, whether it was discontinued for more than 6 months after initiating it, and whether this occurred during the first year of diabetes diagnosis. Sample Adults who reported taking insulin were also asked about insulin-related unmet needs due to cost in the past year, such as, skipping doses, taking less than needed, and delaying buying insulin to save money. Sample Adults who reported having been told by a doctor or health professional that they had diabetes were asked how often they felt overwhelmed by the demands of living with diabetes and how often they felt overwhelmed by the demands of living with diabetes compared with the time before the coronavirus pandemic. Sample adults diagnosed with diabetes during the pandemic were not excluded from being asked the latter question.

NCCDPHP sponsored four questions on epilepsy in the Sample Adult EPI section. Sample Adults were asked whether they had ever been told by a doctor or health professional that they had a seizure disorder or epilepsy, and those diagnosed with the disorder were also asked, whether they were currently taking antiseizure medication, the number of seizures in the past year, and whether they saw a neurologist or epilepsy specialist in the past year. These questions are fielded during 2021-2022 to provide national and subpopulation estimates of epilepsy prevalence. Epilepsy content was fielded in the NHIS prior to the 2019 NHIS redesign.

NCEZID sponsored two questions in the RCN section to provide national estimates for lifetime and current prevalence of chronic fatigue syndrome. Sample adults were asked about whether a doctor or other health professional had told them that they had chronic fatigue syndrome (CFS) or Myalgic Encephalomyelitis (ME), and if they currently have it. There are few national population-based prevalence estimates of CFS/ME. Additionally, the COVID-19 pandemic has increased the need for national prevalence estimates because post-infection CFS/ME symptoms have been reported to be associated with COVID-19 (Bansal, Gubbi and Koch, 2022)). Estimates of lifetime and current CFS/ME can provide data to assist in planning for the care of individuals living with this condition.

Major Recodes

Age of diabetes diagnosis. The age when first told by a doctor or health professional that the Sample Adult had diabetes is top-coded at age 85.

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 of 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, prostrate, 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 = kilograms /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 > = 25 to <30; and obese is BMI > = 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 63–76 inches for men and 59-70 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.

Missed workdays due to asthma. The number of days that Sample Adults with current asthma was unable to work or get work done around the house in the past 12 months because of their asthma was top-coded at 30 or more days.

Number of cancers. A summary recode was created that indicates the number of 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 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). This recode reflects the number of cancers mentioned even when the same kind of cancer was mentioned more than once.

Number of seizures. The response categories for the number of seizures in the past 12 months (0, 1, 2-3, 4-10 and more than 10) was top-coded at "4 or more" seizures, and values 7 (refused), 8 (not ascertained), and 9 (don't know) were retained.

Years since diabetes diagnosis. This recode is the calculation of the Sample Adult's current age (in years and top-coded at 85) minus the age (in years) when first told to have diabetes. This difference may be underestimated for Sample Adults with a top-coded age. The highest number of years since diabetes diagnosis available in the public data release is 84 years. Due to confidentiality concerns, Sample Adults aged 85 and older who were diagnosed at age 85 and over, are coded as 96. This recode updates the recode available in previous years for years since diabetes diagnoses.

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.

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 – 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 #11.

Table 11. Functional limitations included in the Sample Adult module, by topic, section and Washington Group Question Set

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

***Additional questions about hearing are part of the rotating core content, rotating every three years starting in 2021.**

****Questions about fatigue are part of the rotating core content, rotating every two years starting in 2020.**

NOTE: 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.

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. More information on the Washington Group and the question sets may be obtained by request to the WG Secretariat at 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. This recode classifies Sample Adults with disability as those reporting "have a lot of difficulty" or "cannot do at all" for 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 meet 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. In this example, 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: https://www.washingtongroup-disability.com/analysis/analysis-overview/.

A cognitive testing report is available for selected WG questions at https://wwwn.cdc.gov/QBank/Report.aspx?1216. For a report that examines differences in survey responses of disability between the set of disability questions from Short Set on Functioning (WG–SS) and set of disability questions developed for the American Community Survey, see, https://www.cdc.gov/nchs/data/nhsr/nhsr161-508.pdf.

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 ask the Sample Adult whether, because of a physical, mental, or emotional condition, they are limited in the kind or amount of work they can perform, have difficulty doing errands alone, and have difficulty participating in social activities. The degree of difficulty (none, some, a lot, cannot do at all) is captured for the errands and social activities questions. Analysts can analyze each question separately or can combine the questions into an indicator to meet the needs of their analysis.

Rotating Content

Starting in 2021, two additional WG-ES questions on hearing are included every three years in the HEA section. These questions ask about the Sample Adults' level of difficulty hearing a conversation with another person in a quiet room and in a noisier room. The questions are not asked of Sample Adults who responded "cannot do at all" to the difficulty hearing question. Sample adults reporting they 'cannot do at all' to the question about

hearing in a quiet room were not asked about hearing in a nosier room. Analysts can analyze each question separately or can combine the questions into an indicator. These questions can be used as part of the hearing domain in the WG extended set (see Table 11).

An additional question on communication is included every three years in the COM section. Sample Adults were asked whether they use sign language.

Sponsored Content

The Administration for Community Living (ACL) sponsored a question in the ADO section. Sample adults reporting having "a lot of difficulty" or "cannot do at all" to the any one of five functioning questions (walking or climbing stairs, communicating, remembering or concentrating, self-care, or doing errands alone) were asked whether the difficulty began before age 22. The purpose of the question was to assist analysts who wish to identify adults with intellectual or developmental disability.

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. Chronic Pain

Rotating Core

Rotating every other year starting in 2019 in the PAI section of the Sample Adult module are questions about frequency and severity of pain in the past 3 months, and pain impact or interference (how often pain limited their life or work activities and how often their pain had an effect on their family and significant others in the past 3 months). 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). Questions about pain frequency and severity were fielded in 2020 as sponsored content, and pain impact was fielded during July–December 2020 as COVID-19 related content.

Rotating every other year are also questions about the amount of pain experienced in the following body 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 11).

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 12. Similar content is also available for the Sample Child.

Table 12: Annual core content measures of health care access, service use, and affordability of care asked of the Sample Adult, by questionnaire section and reference periods.

Measure	Section	Reference period
Immunizations		
Flu vaccine	IMS	Last 12 months, Month and Year
Pneumonia vaccine	IMS	Ever
Number of pneumonia vaccines received	IMS	Ever
Medical Care		
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
Mental Health Care		
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

Table 12 continued: Annual core content measures of health care access, service use, and affordability of care asked of the Sample Adult, by questionnaire section and reference periods.

Measure	Section	Reference period
Prescription Medication		
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
Problems Paying Medical Bills		
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	Current
***These are family-level replicate questions asked once per family.**		

Sponsored Content

The IMS section includes several questions sponsored by NCIRD regarding CDC recommended vaccinations for flu, shingles, tetanus, hepatitis A, hepatitis B, and starting in April 2021 (quarter 2) COVID-19 (See Table 13).

Table 13. Sponsored content about immunization measures asked of Sample Adults, by population, reference period, and additional information collected: 2021 NHIS.

Vaccine	Population	Reference Period	Additional information collected
COVID-19*	Adults 18 years and over	Ever	Number of vaccines Month and year of most recent and next most recent vaccination Additional dose needed
Flu	Female Adults aged 18–49 or age unknown	Before/during/after a current or recent pregnancy (see flu vaccination timing, below)	Not applicable
Hepatitis A	Adults 18 years and over	Ever	Not applicable
Hepatitis B	Adults 18 years and over	Ever	Not applicable
Shingles	Adults aged 50 and over	Ever, when last	Type Number of Shingrix vaccines
Tetanus	Female Adults aged 18–49 or age unknown who had a live birth in the past 12 months	During pregnancy	Not applicable
***Similar content was collected about Sample Children. Questions about COVID-19 vaccination began fielding in Quarter 2 (April) among adults aged 18 and over, and in Quarter 3 (July) among children aged 12-17 years.**			

While overall content regarding shingles vaccines did not change in 2021, there were slight changes to the wording of the questions about whether Zostavax or Shingrix vaccines were received. Variable names have been changed to reflect these wording changes.

In December 2020, the FDA issued an emergency use authorization for the first COVID-19 vaccine, administered as two doses three weeks apart, for people aged 16 and older (FDA, 2020). The Sample Adult was asked whether they had received a COVID-19 vaccination, the number of vaccinations they received, the dates (month and year) of the most recent and next most recent vaccinations, and whether a health professional indicated that they needed an additional vaccine dose.

NCIRD sponsored questions about hepatitis exposure, including international travel to areas of the world with high endemicity of hepatitis A and B infections, and whether the Sample Adult ever lived with someone who had hepatitis.

NCIRD also 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, and two questions for Sample Adults aged 18 and over about work or volunteer activities in health care settings. The purpose of these questions is to permit analysis of vaccine uptake by pregnant women and health care workers.

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 through July, were asked about pregnancy status during August through March as follows: since August 1st of last year if interviewed between January through March; from August of last year through March of current year if interviewed April through July; and since August 1 of current year if interviewed August through December. Those who reported to be currently pregnant or had a recent pregnancy during August through March, and who had received a flu vaccine in the past 12 months were asked: "Earlier you said you were pregnant sometime [since August 1st, {prior year}/from August {prior year} through March {current year}/since August 1st, {current year}] Did you get a flu vaccination before, during, or after your pregnancy?"

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® vaccine, and 2017 for the year of the most recent Shingrix® vaccine. Sample Adults reporting years earlier than these lower limits were assigned the value 9996.

Month and year of two most recent COVID-19 vaccinations. Recodes were created to update observations with inconsistent timing of vaccination in relation to when the vaccine became available and address apparent respondent error or data entry errors in reporting the order of doses or when combined with timing of interview. Less than 5% of original vaccination dates were updated. These recodes correct for possible errors in recall or data entry and prevent potential disclosure of individuals who may have participated in vaccine research. For adults whose reported vaccination year was 2020 and vaccination month was January through August, the year of vaccination was changed to 2021. For those whose reported vaccination year was 2020 and vaccination month was September through November, the vaccination date was changed to December 2020 for

the first dose, and January 2021 for the second dose if two doses were reported, —to reflect when vaccines became officially available. These recodes also switch the vaccination dates if the most recent vaccination date reported was before the next most recent vaccination.

V. Preventive Care and Services

Rotating Core

Aspirin Use. Rotating every other year starting in 2019 are questions about aspirin use. In the ASP section, Sample Adults aged 40 and over were asked if a doctor or other health professional ever told them to take a low-dose aspirin each day to prevent or control heart disease, whether they were following that advice; and if a doctor or other health professional had advised them to stop taking it. Those who had never been advised to take a low-dose aspirin, were asked whether they were taking a low-dose aspirin each day, on their own, to prevent or control heart disease. In April 2022, the United States Preventive Services Task Force (USPSTF) updated the 2016 recommendation on the use of low-dose aspirin for the prevention of cardiovascular disease (CVD) (USPSTF, 2022). The USPSTF concluded that low-dose aspirin has a small net benefit in reducing the risk for cardiovascular events in adults ages 40 to 59 years who have a 10% or greater 10-year CVD risk, and there is no net benefit to initiating aspirin use for the primary prevention of CVD events in adults aged 60 years or older (USPST, 2022). Due to a programming error, questions about aspirin use were not fielded in Quarter 1 and are only available for Quarters 2-4 (April–December) of 2021.

Table 14. Rotating content about cardiovascular and cancer screening measures asked of Sample Adults, by population asked and period measured.

Topic	Population asked	Period measured
Cardiovascular checks		
Arterial blood pressure	Adults 18 years and over	When last
Blood cholesterol	Adults 18 years and over	When last
Blood Glucose	Adults 18 years and over	When last
Breast cancer screening		
Mammogram	Females 30 years and over	Ever; When last
Cervical cancer screening		
Pap smear or Human Papillomavirus (HPV) test	Females 18 years and over	Ever; When last
Hysterectomy*	Females 18 years and over	Ever
Colorectal cancer screening		
Colonoscopy**	Adults 40 years and over	Ever; When last
Sigmoidoscopy**	Adults 40 years and over	Ever; When last
CT colonography/virtual colonoscopy	Adults 40 years and over	Ever; When last
Fecal immunochemical test (FIT)/fecal occult blood test	Adults 40 years and over	Ever; When last

***The response "yes" to the question on hysterectomy has been edited to include women who reported in a previous question having had a hysterectomy as a reason for never having or not having had in the last 5 years a test for cervical cancer screening.**

****Sample adults who had a colonoscopy or sigmoidoscopy but did not know which one were only asked for the most recent time of this exam.**

Preventive Services. Rotating every other year in the PRV section of the Sample Adult module is content about clinical checks of blood glucose, cholesterol, arterial blood pressure, and screenings for colorectal cancer, and breast and cervical cancer among women, see Table 14. Questions about cardiovascular checks were asked of all adults and questions about cancer screening were asked among adults who reported their age, and their sex for sex-specific cancer check-ups.
The USPSTF recommends screening for colorectal cancer screening for adults aged 45-75 years, and selective screening for adults aged 76 to 85 years (Lin et al., 2021). There are multiple factors that influence which tests clinicians and patients may use —frequencies of screenings, home-based or clinical setting, stool-based or direct visualization, preparation procedures, sedation during the test, and follow-up procedures for abnormal findings.

Analysts comparing selected colorectal test measures between 2021 and 2019 NHIS should note a change in the population asked these questions. In 2021, questions about history of CT colonography or virtual colonoscopy and of fecal immunochemical test (FIT) or fecal occult blood test were asked of all Sample Adults aged 40 years and over, however, in 2019, these questions were restricted to those aged 40 and over who reported any other kind of cancer screening test after being asked about colonoscopy or sigmoidoscopy. Questions were updated in 2021 to improve measurement of colorectal testing options among adults.

The USPSTF recommends screening women aged 21 to 65 years for cervical cancer every three years with pap smear test (cervical cytology) alone, and in women aged 30 to 65 years every five years with high-risk human papillomavirus (hrHPV) testing alone or dual testing of hrHPV and pap smear (Melnikow et al., 2018.). The USPSTF also recommends mammography screening for breast cancer every two years in women aged 50 to 74 years (Siu, 2016).

Sponsored Content

NCCDPHP sponsored two questions asked in the PRV section that asked Sample Adults with diabetes about the last time they had their hemoglobin A1C level checked by a doctor, nurse, or other health professional, and the number of times monitored if checked in the past 12 months.

NCCDPHP and NCI sponsored several questions about breast, cervical, colorectal, and prostate cancer screenings in the PRV section, see Table 15 for question topic and which Sample Adults were asked about them. Some of the listed cancer screening content was previously fielded in the NHIS in 2019 or in prior years. Analysts comparing measures about cancer screening between 2019 and 2021 should note variable name updates which are indicative of changes in universe (i.e., who was asked the question), response options or question text. For example, questions about history of Cologuard testing were asked of all Sample Adults 40 and over in 2021, but in 2019 these were only asked of Sample Adults 40 and over who previously reported having any other (than a colonoscopy or sigmoidoscopy) cancer screening, and who had a FIT test. Similarly, questions about having a doctor or health professional recommend colorectal screening in the past 12 months have also had their universe updated in 2021 to better account for the colorectal testing options.

Major Recodes

Number of A1Cs. Responses 5 to 95 times, and "96 or more" for the number of A1C tests in the past 12 months were top-coded as “5 or more,” and responses corresponding to "refused," "not ascertained," and "don’t know" were retained with values 7–9, respectively.’’

Table 15. Sponsored content on cancer screening, by population and reference period: 2021 NHIS.

Topic	Population asked	Reference period
Breast cancer screening		
Age of first mammogram	Females 30 years and over	N/A
Mammogram payment/coverage	Females 30 years and over who ever had a mammogram	At most recent screening
Reason for testing	Females 30 years and over who ever had a mammogram	At most recent screening
Reason for not testing	Females 30 years and over who never had a mammogram or had a mammogram more than 2 years ago	N/A
Cervical cancer screening		
Pap test	Females 18 years and over	At most recent screening
HPV test	Females 18 years and over	At most recent screening
Told which cervical test done	Females 18 years and over who had a cervical test	N/A
Follow-up for cancer/pre-cancerous cells	Female 18 and over who had a cervical cancer test within the past 5 years	N/A
Reason for testing	Females 18 years and over who ever had a cervical test	At most recent screening
Reason for not testing	Females 18 years and over who never had a cervical test or last test was more than 5 years ago	N/A
Colorectal cancer screening		
Cologuard*	Adults 40 years and over	Ever; When last
Colonoscopy test reason	Adults 40 years and over who ever had a coloscopy	At most recent screening
Colonoscopy payment/coverage	Adults 40 years and over who ever had a coloscopy	At most recent screening
Doctor/health professional recommended colorectal screening	Adults 40 and over never tested for colorectal cancer, or whose last test exceeded the recommended time interval between testing, or is unknown**	Past 12 months
Type of tests recommended by doctor/health professional	Adults 40 and over who had a colorectal cancer test recommended	Past 12 months
Prostate cancer screening		
Prostate-specific antigen (PSA) test	Males 40 years and over	Ever; When last
Reason for PSA	Males 40 years and over who ever had a PSA test	At most recent screening
Who suggested PSA test	Males 40 years and over who ever had a PSA test	N/A

N/A is not applicable.
***Sample adults who reported "yes" to having had a Cologuard test were also asked whether this was part of the Fecal immunochemical test (FIT) reported in an earlier question, and only those reporting "no," "don't know," or "refused" were asked for the last time they had a Cologuard test.**
****See the universe description in the Questionnaire or Sample Adult codebook for more information.**

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 someday 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 also asked about the use of electronic cigarettes or other electronic vaping products, even one time in their entire life, and if so, 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 every day smoker; current someday 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 2021 in the SPD section is the Kessler 6 (K6) nonspecific distress scale (Kessler, 2003). The six-question K6 was developed to identify persons with a high likelihood of having a diagnosable mental illness and associated functional limitations, using as few questions as possible. The K6 scale was designed to be sensitive around the threshold for the clinically significant range of the distribution of nonspecific distress to maximize the ability to discriminate cases of serious mental illness (SMI) from non-cases. Since its development, several studies have supported its validity (Kessler 2002, Kessler 2010, Furukawa 2003).

The set of questions in the K6 asks Sample Adults to assess the frequency of feeling sad, nervous, restless, hopeless, that everything was an effort, and worthless, in the past 30 days. The response option for each question is scored as follows: 0 for "none of the time" 1 for "a little of the time," 2 for "some of the time," 3 for "most of the time," and 4 for "all of the time." The points for each question are summed to produce a total score between 0 and 24. Total scores may be used in analysis, and a pre-determined cut-off value of the total score is available for assessing presence of serious psychological distress (see major recodes, below) (Kessler, 2003). Note that the scoring for each question for the presence of serious psychological distress ranges 0-4 and is in reversed order from the actual response values used in the data and documentation which range from 1 "all of the time" to 5 "none of the time."

Sponsored Content

Life satisfaction is an aspect of general well-being that represents people's subjective assessments of their lives overall and reflects factors such as health, work, and income, among others (OECD, 2013). The American Psychological Association defines life satisfaction as "the extent to which a person finds life rich, meaningful, full, or of high quality" (APA, 2022). Measures of life satisfaction, including the items used in NHIS, have been shown to be valid and reliable (Cheung and Lucas, 2014; OECD, 2013). Life satisfaction will be used as a Healthy People 2030 Overall Health and Well-being Measure to estimate and track adult well-being at the national level (Ochiai et al., 2021).

NCCDPHP and ODP sponsored two life satisfaction questions for Sample Adults. One asked the respondent to rate how they felt about their life as a whole these days using a scale of 0 to 10, where 0 means "very dissatisfied" and 10 means "very satisfied," and the second question asked about how satisfied they were with their life using response options of very satisfied, satisfied, dissatisfied, and very dissatisfied. Each life satisfaction question was administered in separate locations of the questionnaire in the sections LS1 and LS2, but the Codebook and Summary documents describe these variables under a new section called LS3. Beginning in quarter 2, April 2021, an experiment was incorporated into the survey to rotate the order in which these two life satisfaction questions were administered. The purpose of this experiment was to determine whether the life satisfaction questions were influenced by contextual effects due to item placement and whether one question was more prone to these effects. One half of the sample received the 11-point question first and towards the beginning of the health content in the Sample Adult module (after health status), and the 4-category question second and towards the end of the health content (after social support). The remaining half of the sample

received the 4-category question first and towards the beginning of the health content, and the 11-point question second and towards the end of the health content. Recodes have been created to combine responses from the same question administered early versus later in the survey (see major recodes, below). A variable identifying the rotation order is available in the file.

Major Recodes

Serious psychological distress. A dichotomous recode was created based on the sum of the responses to the six K6 questions (Kessler 2003), where scores 13 or more represent 1) experiencing serious psychological distress, and scores 0—12 represent 2) not experiencing serious psychological distress. Adults who responded, "don't know," "refuse" or whose responses to any of the six psychological distress questions was not ascertained, are excluded from the calculation, and coded as 8 (not ascertained).

Life satisfaction. Recodes were created for the 11-category question and the 4-category question. Regardless of when it was asked in the interview, responses based on the same life satisfaction question were combined into a single recode.

VIII. Injuries

Rotating Core

For two consecutive years every three years starting in 2020, the Sample Adult module included questions about repetitive strain injuries (RSI) in the REP section, followed by questions about sudden injuries in the INJ section. RSIs are defined as injuries caused by repeating the same movement over an extended period. Examples of RSI include carpal tunnel syndrome, tennis elbow, or tendonitis. Sudden injury may occur accidently or on purpose, and it may be self-inflicted or caused by others. All questions in the REP and INJ sections were asked with the reference period of past 3 months.

The REP section asked whether the Sample Adult had an RSI, and if the RSI had limited usual activities for at least 24 hours. Those limited in their usual activities were asked about visiting a doctor or other health professional for the RSI and being told by a doctor or other health professional that any of their RSIs were work-related. Sample Adults limited in their usual activities due to the RSI were also asked about the impact on their work: number of workdays missed due to the RSI, expecting to miss additional workdays due to the RSI, stopping or changing job due to the RSI, and changing work activities due to the RSI. Since the employment questions are asked later in the interview and those questions are in reference to the past the week, interviewers were instructed to ask the Sample Adult reporting zero missed workdays if they had worked at all in the past 3 months. Those who did not work in the past 3 months were included in a separate category ('91') for number of missed workdays, but they were included in the 'no' response category for questions about changing job or job duties.

The INJ section asked whether the Sample Adult had a sudden injury, and if the sudden injury had limited usual activities for at least 24 hours after the injury occurred, also defined as significant injuries. Those with a sudden significant injury that limited their usual activities were asked about the number of significant injuries, place where they were injured (e.g., at home or at work), activity they were doing when they were injured (e.g., playing sports, driving, or doing household chores), medical care they received (e.g., saw a doctor, visited ER, or

hospitalized overnight), and impact on work (e.g., number of workdays missed, expected workdays missed, or stopped or changed job). Also refer to Table 16 for measures of sudden injury.

Table 16. Rotating core content on adult sudden injuries: 2020–2021 NHIS
Measure
Any injury
Number of significant injuries
Place where injured
Home
Work
Activity when injured
Sports
Household chores
Motor vehicle accident: driver, passenger, bicyclist, pedestrian, or something else
Type of Injury
Fall
Motor vehicle accident
Medical care received
Saw a doctor or other health professional
Visited Emergency Room
Hospitalized overnight
Impact on work
Number of workdays missed
Expected workdays missed
Stopped or changed job
Broken bones due to injury
Stitches or staples due to injury

Major Recodes

Missed workdays due to RSI. Number of workdays missed due to repetitive strain injury was top-coded at 10 or more days, and values 91 (did not work), 97 (refused), 98 (not ascertained), and 99 (don't know) were retained.

Missed workdays due to sudden injury. Number of workdays missed due to sudden injury were top-coded at 10 or more days, and values 91 (did not work), 97 (refused), 98 (not ascertained), and 99 (don't know) were retained.

Significant injuries. Number of significant injuries was top-coded at 10 or more days, and values 97 (refused), 98 (not ascertained), and 99 (don't know) were retained.

IX. COVID-19

In March 2020, the World Health Organization declared a global pandemic of coronavirus disease 2019 (COVID-19), caused by a severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). A national emergency was declared in the United States, and states began issuing stay-at-home orders, mask mandates, and capacity limits at businesses to slow the rate of new infections. Social distancing measures were also recommended for shared spaces including those in the workplace. The virus was first detected in December 2019 in Wuhan, China, and in January 2020, CDC had confirmed the first case of COVID-19 in the United States. In the United States, testing for COVID-19 became available in February 2020, and the first COVID-19 vaccine became available in December 2020.

Emerging Content

Since July 2020, several questions have been added in the NHIS in response to the COVID-19 pandemic, see Table 17. Sample Adults were asked in the CVD section about positive COVID-19 diagnosis, testing and perceived symptom severity. Sample Adults were asked in the ACC section about unmet medical needs due to the pandemic, receipt of telemedicine in the past year and due to the pandemic. The CGR section included questions about receipt of at-home care from a from a nurse or other health professional and friend or family member, unmet need for at-home care due to the pandemic, and whether a friend or family member provided some or all of at-home care not provided by a nurse or health professional due to the pandemic. Two questions in the SOS section asked the Sample Adult's perception of how often they receive the social and emotional support they need and the receipt of this support in comparison to the past 12 months.

Sponsored Content

Since July 2020, sponsored content on cancer care (CNV section) and social distancing in the workplace (SDW section) were added in response to the COVID-19 pandemic. NCCDPHP and NCI sponsored content about the receipt of medically indicated treatment and medical care for cancer and whether these were unmet due to the pandemic. NIOSH sponsored content about social distancing measures in effect to help keep people apart and the need to work closer than six feet from other people at their workplace, among Sample Adults who were working in the past week and those who did not work in the past week but worked in the past year. The variable names for these questions and recodes were updated in 2021 due to a change in universe. In 2021, these questions were asked of Sample Adults who performed seasonal or contract work in the past 12 months instead of all seasonal or contract workers.

In 2021, NIDCD sponsored content about new loss of smell or taste due to COVID-19 in the TSC section. Sample Adults who reported they were diagnosed with COVID-19 or had a positive COVID-19 test were asked about losing their sense of smell or smelling odors that were not there, and about losing their ability to taste or having unwanted tastes or sensations, and whether they had fully, partially, or had not recovered their sense of smell and taste.

Table 17. COVID-19 measures

Topic (Section)	Reference period
Positive COVID-19 Diagnosis (CVD)*	
Diagnosed	Ever
Tested	Ever
Positive test	Ever
Symptom severity	Symptoms at their worst
Taste and Smell – COVID (TSC) [1]	
Loss of smell	During COVID-19 disease
Recovery of smell	Current
Loss of taste	During COVID-19 disease
Recovery of taste	Current
Access to Care (ACC)*	
Delayed medical care	During pandemic
Did not get needed medical care	During pandemic
Telemedicine	Past 12 months; During pandemic
Cancer Care (CNV)[2]	
In treatment or medically indicated	During pandemic
Needed other cancer medical care	During pandemic
Treatment changed, delayed, or cancelled	During pandemic
Other cancer medical care changed, delayed, or cancelled	During pandemic
Caregiving Received (CGR)	
Received care at home from a nurse or other health professional	Past 12 months
Did not get needed care at home from a nurse or other health professional	During pandemic
Received care at home from a friend or family member	Past 12 months
Did not get needed care at home from a friend or family member	During pandemic
Friend or family member provided care at home not provided by nurse or other health professional	During pandemic
Social Support (SOS)	
Social and emotional support	Current; Compared with 12 months ago
Social distancing at work (SDW)[3]	
Social distancing measures in effect	Current; Since the start of the pandemic (working at main job or business); Ever (not working but held a job in the past 12 months)
Work closer than 6 feet to other people	Current (working at main job or business); At last job (not working but held a job in the past 12 months); When social distancing measures in effect/not in effect (working at main job or business, and not working but held a job in the past 12 months)

***These measures are also available for Sample Children.**
[1] Sponsored content from NIDCD.
[2] Sponsored content from NCCDPHP and NCI.
[3] Sponsored content from NIOSH.

Note that additional measures fielded among Sample Adults in response to the COVID-19 pandemic are found in the IMS, ISN, RCN, and JOB sections. See descriptions for COVID-19 vaccination (IMS section) under "Health Care Access and Health Service Utilization, additional health conditions (ISN and RCN sections) under "Health Status and Conditions," and missed workdays in the past 3 months due to COVID-19 infection (JOB section) under "Characteristics about the Sample Adult and Sample Child."

Major Recodes

Social distancing at any time (JOBANYSD1_A)**.** This recode summarizes a set of questions into whether there were ever any social distancing measures at the Sample Adult's workplace, among those who were currently working (including those on temporary leave last week, those with seasonal/contract work in the past 12 months, and those who work but not for pay) and those who worked anytime in the past 12 months. The recode is based on responses to the following questions: whether there are currently social distancing measures at the Sample Adult's main job; whether the Sample Adult's main job implemented social distancing measures at any time since the COVID-19 pandemic began; and whether those who are not currently working but worked in the last 12 months ever had social distancing measures in place at their main job when they were still working there. For Sample Adults who are working but do not have social distancing measures currently in place or who don't know or refused to answer, final responses to this recode are based on the question on whether social distancing measures were ever instituted, regardless of their response about current social distancing measures.

Frequency of working close to others when social distancing measures were in effect (CLSWRKSD1_A). Among Sample Adults who had social distancing measures in effect at their main job at any time, this recode summarizes the frequency that Sample Adults work(ed) closer than 6 feet to others when social distancing measures were in effect. It is based on the response to one of three questions about frequency of working closer than 6 feet to others: WRKCLSSD1_A for those who currently have social distancing measures at work; SDMSRSOFT1_A for those who do not currently have social distancing measures at work but whose jobs did at some point since the coronavirus pandemic began; and RJCLSSD1_A for those who are no longer working but had social distancing measures in place when they were working in the past 12 months. Response options are "all of the time," "some of the time," "most of the time," and "none of the time." An additional response option was added to categorize those who only worked at their main job when the social distancing measures were not in effect.

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 2–17 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.

Rotating Core

Rotating every three years starting in 2021 is content about respiratory, food, and skin allergies in the ALG section of the Sample Child module. The Sample Child respondent was asked whether the child experiences hay fever, seasonal or year-round allergies, an allergy to one or more foods, and an itchy rash due to eczema or atopic dermatitis. For those children experiencing allergy symptoms, the Sample Child respondent was asked whether a doctor or other health care professional had ever told them that the child has these respective allergies. Similar content is also available for the Sample Adult.

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 compose the 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 about children 2–4 years of age (CFM 2–4) and questions about 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.

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. For a list of questions asked in each set, and the associated functioning domain, questionnaire section, and respective age range, see Table 18. In addition, several of the domains have questions to provide information on the use of accommodations. Questions about the use of equipment or assistance with walking were asked about children who had difficulty walking both with and without equipment or assistance. More information may be obtained by request to the WG Secretariat at 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. Consult the WG website for guidance on the creation of these indicators. Changing the threshold for either the number of domains the respondent identifies the child having difficulty with or the child's degree of difficulty can create different identifiers that will capture different populations. For example, a recode that includes Sample Children with the response "some difficulty" to any of the domains will capture a greater proportion of the population than a recode that includes only Sample Children with the response "cannot do at all" to any of the domains. In this example, 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.

Table 18. Annual core content measures of functional limitations measured in the Sample Child module, by domain, question topic, questionnaire section and age range of question

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	COM	2–4
Communication	Difficulty being understood by people inside of 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

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. This recode classifies children with disability as those with the responses “a lot of difficulty” or “cannot do at all” for at least one of the questions asking about the Sample Child’s difficulty seeing, hearing, walking, dexterity, communication, learning, and playing, or with the response “cannot do at all” to the question about controlling behavior. The remaining Sample Children, with the responses "some difficulty" or “no difficulty" to at least one question (and who do not have the responses “a lot of difficulty” or “cannot do at all” for any of the questions) are classified as without disability. Sample Children with a response of 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. This recode classifies children with disability as those with the responses “a lot of difficulty” or “cannot do at all” for at least one of the questions asking about the Sample Child’s difficulty seeing, hearing, walking, self-care, communication, learning, remembering, concentrating, accepting change, controlling behavior, making friends or the response “daily” to questions asking how often the Sample Child feels anxious, nervous, or worried or feels depressed. The remaining Sample Children, with the responses "some difficulty" or “no difficulty" to at least one question (and who do not have the responses “a lot of difficulty” or “cannot do at all” or “daily” for any of the questions) are classified as without disability. Sample children with a response of "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 19.

Table 19. Annual core content measures of health care access, service use, and affordability of care asked of the Sample Child, by questionnaire section and reference periods

Measure	Section	Reference period
Immunizations		
Flu vaccine (any, number of vaccines up to 2)	IMM	Last 12 months, Month and Year
Medical Care		
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
Prescription Medication		
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
Problems Paying Medical Bills		
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	Current
***These are family-level replicate questions asked once per family.**		

Rotating Core

In 2021, questions on receipt of select mental health services were asked of the Sample Child (see Table 20). These questions are part of a set of NHIS rotating content with a focus on health service utilization that were retained in 2021 to assess the impact of COVID-19, as well as to provide compatible measures with an ongoing special project on adolescents whose parents completed the NHIS Sample Child interview.

Table 20. Additional measures of health care access, by questionnaire section and reference period: 2021 NHIS

Measure	Section	Reference period
Mental Health Care		
Took medication for emotions/mental health	MHC	Last 12 months
Received therapy or counseling from mental health professional	MHC	Last 12 months

Emerging Content

Questions were added in 2021 to the Sample Child UTZ section which asked whether Sample Children aged 12-17 had time during their last medical visit or their last wellness visit to speak to the health care professional privately. An additional question asked if there was one or more than one person that the Sample Child respondent thought of as the Sample Child's personal health care professional. These questions assess whether the adolescent may have received services or advice that the parent or guardian may not be aware of, and whether adolescents had the opportunity to discuss health care with their provider as part of a transition into adulthood.

Sponsored Content

Beginning in July 2021, questions about COVID-19 vaccination sponsored by NCIRD were added to the Sample Child module in the CVV section for children aged 12-17 years. In December 2020, the FDA issued an emergency use authorization for the first COVID-19 vaccine, administered as two doses three weeks apart, for people aged 16 and older (FDA, 2020). This emergency use authorization was expanded to include children aged 12-15 years in May of 2021 (FDA, 2021). The Sample Child respondent was asked whether the child has had a COVID-19 vaccination, the number of vaccinations the child has received, the dates (month and year) of the most recent and next most recent vaccinations, and whether a health professional indicated that the child needs an additional vaccine dose. Similar content was collected from Sample Adults starting in January 2021.

Major Recodes

Month and year of two most recent COVID-19 vaccinations. Recodes were created to update observations with inconsistent timing of vaccination in relation to when the vaccine became available for the child's age and address apparent respondent error or data entry errors in reporting the order of doses or when combined with timing of interview. Less than 10% of original vaccination dates were updated. These recodes correct for possible errors in recall or data entry and prevent potential disclosure of individuals who may have participated in vaccine research. For children aged 16-17 whose reported vaccination year was 2020 and vaccination month

was January through August, the year of vaccination was changed to 2021. For those whose reported vaccination year was 2020 and vaccination month was September through November, the vaccination date was changed to December 2020 for the first dose, and January 2021 for the second dose if two doses were reported, —to reflect when vaccines became officially available. For children aged 12-15, vaccination dates reported in 2020 were changed to the year 2021, and vaccination dates in 2021 with a reported month before May—when vaccines became available for children aged 12-15—were changed to May 2021. These recodes also switch the vaccination dates if the most recent vaccination date reported was before the next most recent vaccination.

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

In the BSC section of the Sample Child module, 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 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.

Emerging Content

Social Support. Beginning in 2021 receipt of social support was identified as an emerging topic for inclusion on the Sample Child module. Presence of a positive adult mentor in a child or adolescent's life is associated with the ability to flourish, complete high school, attend college, reduced problem behaviors (e.g., gang membership), improved psychological well-being and health (DuBois et al., 2005). Two questions related to the frequency of receiving social and emotional support and having an adult in their life that makes a difference were included in the SOS section. Receipt of social and emotional support was derived from the Sample Adult questionnaire, while the adult that makes a difference was derived from the National Survey of Children's Health. Social support questions were only asked among adolescents, or Sample Children aged 12-17.

Bullying. Beginning in 2021 bullying was identified as an emerging topic for inclusion on the Sample Child module. Bullying victimization is associated with negative health outcomes including increased suicide attempts, depressive symptoms and lower school performance and school attachment (Schneider et al., 2012). Three questions about the frequency of being bullied, recent electronic bullying and frequency of bully perpetration

were asked in the BLY section. Bullying questions were only asked among adolescents, or Sample Children aged 12-17. Bullying measures are adapted from the Youth Risk Behavior Surveillance System (YRBSS).

Sponsored Content

NCCDPHP and ODP sponsored a question on satisfaction with life asked about Sample Children aged 12-17 in the LS1 section. The Sample Child respondent was asked to rate how they thought the child felt about their life as a whole these days using a scale of 0 to 10, where 0 means "very dissatisfied" and 10 means "very satisfied." This same measure was also collected from Sample Adults.

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). Rotating every other year are four questions in the SLE section that asked the Sample Child respondent 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, and had a parent who was incarcerated after the child was born. These questions were adapted from the Adverse Childhood Experience (ACE) questions that have previously been fielded in the National Survey of Children's Health (NSCH, 2021).

In 2021, four additional questions were included in the SLE section as either emerging or sponsored content. The combined set of eight questions are intended to create a more comprehensive picture of potential stressful life events during the childhood years while providing compatible measures for an ongoing online survey of adolescents whose parents completed the NHIS Sample Child interview.

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

Emerging Content

The Sample Child respondent was asked whether the child had ever been treated or judge unfairly because of race or ethnic group, and among Sample Children aged 12-17 years, whether the child had been treated or judged unfairly because of sexual orientation or gender identity. Both measures were also asked on the NSCH beginning in 2020.

Sponsored Content

NCIPC sponsored two stressful life event questions. The Sample Child respondent was asked whether the child had ever lived with a parent or adult who was verbally abusive, and whether the child ever experienced unmet

basic needs such as not having enough to eat, not going to a doctor when was sick, or not having a safe place to stay. These questions were adapted from the Optional Questions List for YRBSS and the ACE Module in the Behavioral Risk Factor Surveillance System.

VI. Injuries

Rotating Core

For two consecutive years every three years starting in 2020, the Sample Child INJ section included questions about sudden injuries in reference to the past 3 months. An introductory statement preceded these questions that told the Sample Child respondent that the questions were about all types of injuries, and that injuries may occur accidently or on purpose, and it may be self-inflicted or caused by others. Questions ask whether the Sample Child had an injury, and if the injury was significant enough that it limited Sample Child's usual activities for at least 24 hours. Additional information about significant injuries included the number of significant injuries, place where the injury occurred (e.g., at home, day care if 2 years old or younger, day care or school if 3 to 5 years old, or school if 6–17 years old), activity they were doing when they were injured (e.g., playing sports or exercising, or an activity related to motor vehicle accident), medical care received (e.g., saw a doctor, visited emergency room, or hospitalized overnight), and impact on daily life (e.g., number of daycare or school days missed and additional days expected to miss). Refer to Table 21 for a list of types of injury measures collected.

Table 21. Rotating core content on child injuries: NHIS 2020–2021

Measure
Any injury
Number of significant injuries
Place where injured
Home
School or daycare
Activity when injured
Sports
Motor vehicle accident: driver*, passenger, bicyclist, or something else
Type of Injury
Fall
Motor vehicle accident
Medical care received
Saw a doctor or other health professional
Visited Emergency room
Hospitalized overnight
Impact on school
Number of school or daycare days missed
Expect to miss additional school or daycare
Broken bones due to injury
Stitches or staples due to injury

Note: The option of driver is unavailable for children under 6 years old.

Emerging Content

Content about head injuries or concussions experienced by Sample Children was added in the TBI section as emerging content during 2020–2021. The Sample Child respondent was asked to think about all head injuries, for example, from playing sports, car accidents, falls, or being hit by something or someone. The questions ask about head injuries that may have occurred anytime in Sample Child's life, ever experiencing concussion symptoms as a result of that head injury (e.g., gap in memory, headaches, vomiting, blurred vision, or changes in mood or behavior), receiving assessment for a concussion from a health care professional, and a diagnosis of a concussion from a health care professional.

Major Recodes

Missed school or daycare days. Number of school days missed due to injury was top-coded at 10 or more days, and values 97 (refused), 98 (not ascertained), and 99 (don't know) were retained.

VII. COVID-19

In March 2020, the World Health Organization declared a global pandemic of coronavirus disease 2019 (COVID-19), caused by a severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). A national emergency was declared in the United States, and states began issuing stay-at-home orders, mask mandates, and capacity limits at businesses to slow the rate of new infections. Social distancing measures were also recommended for shared spaces including those in the workplace. The virus was first detected in December 2019 in Wuhan, China, and in January 2020, CDC had confirmed the first case of COVID-19 in the United States. In the United States, testing for COVID-19 became available in February 2020, and the first COVID-19 vaccine became available in December 2020.

Emerging Content

Beginning in July 2020 (quarter 3), several questions were added to the NHIS survey in response to the COVID-19 pandemic and continued through 2021. The Sample Child module included questions in the CVD section about positive COVID-19 diagnosis, testing and perceived symptom severity. The Sample Child module also included questions in the ACC section about the Sample Child's unmet medical needs due to the pandemic, receipt of telemedicine in the past year, and receipt of telemedicine due to the pandemic.

Questions in the CVD and ACC sections were asked about Sample Children aged 0 to 17 years, and similar content is also available for Sample Adults. Measures on COVID-19 vaccination were added for Sample Children aged 12– 17 years in July of 2021 (quarter 3). See descriptions for the CVV section under "Health Care Access and Health Service Utilization."

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 starting in 2019 and continuing through 2021 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 types of coverages, 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 vison 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 measured included the following: 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 measured included the following: 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 a family 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. It is important to note that due to confidentiality concerns the HIKIND03_C is not available on the Sample Child file. Responses to HIKIND03_C have been combined with responses to HIKIND01_C and are made available in HIKIND01R_C.

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 22 below, for specific types of healthcare coverage and noncoverage because of the complicated editing process that takes place in the INS sections.

Table 22. Annual core content of health insurance recode variables for the Sample Adult and Sample Child

Type of health insurance coverage	Sample Adult file	Sample Child file
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 NHIS collects information across different modules and sections that describe the Sample Adult and Sample Child's sex, age, Hispanic origin, and race. Table 23 lists the variables that summarized the final public-use variables on sex, age, Hispanic origin, and race available for them. These variables are described in the HHC section of the Codebooks.

Table 23. Annual core content on public-use variables for sex, age, Hispanic origin and race variables

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

***During rostering, household respondents who refused or don't know the age of a household member are then asked a set of age range questions about the household member. The public-use data file includes information from the household respondent on whether the person selected to be the Sample Adult is under 65 years or age or 65 or older (AGE65). In the Sample Adult and Sample Child modules, the respondent can then provide the missing age, correct it, or also decline to provide this information. AGEP_A and AGEP_C are based on the final age information.**

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 the four mentioned or reporting more than one race, other than including AIAN, were combined into the "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 of the Sample Adults and Sample Child modules. To protect confidentiality, information on U.S. state or territory of birth and detailed citizenship status is 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 their 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 EDUCP_A describes the highest educational attainment of the Sample Adult, top-coded to combine professional school degree (e.g., MD, DDS, DVM, JD) with doctoral degree (e.g., PhD, EdD). Additionally, similar 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"). Persons responding for the Sample Children are not asked about highest school grade or education the Sample Child has 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 for pay, when they last worked for pay, and among those who worked for pay in the past year, usual number of hours worked, work benefits, and days of sick leave taken. The variable names for employment questions were revised in 2021 due to changes in question order and universe.

Beginning in 2021, the annual content of the EMP section fielded during 2019-2020 was restructured to ask for the reason for not working for pay last week and the last time the Sample Adult worked for pay earlier in the set of questions. This change informed the universe of subsequent questions. For example, Sample Adults who did seasonal or contract work but did not work for pay last week were now asked about the last time they worked for pay (and like in 2019-2020, it was also asked of those who did not work last week because they were retired, unable to work for health reasons, taking care of the house/family, going to school, or for some other reason, or refused or answered don't know). Sample Adults who worked anytime in the past year (for pay or not for pay at family business) were now asked about typical total number of hours worked at all job or businesses, and whether they work 35 or more hours was now only asked as a follow-up question to those who refused or didn't know how many hours they usually work or worked (both questions previously asked of those who only worked for pay last week, and hours worked were unknown or less than 35 hours). Lastly, Sample Adults who worked anytime in the past year were now asked about job-sponsored health coverage, paid sick leave, and workdays missed due to illness or injury (previously asked of those working at or on temporary leave last week, performed seasonal/contract work regardless of when they last worked for pay, and those who work, but not for pay at a family business).

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, or family or maternity leave. Those who were working or temporarily absent from work were then asked how many hours in total they usually work at all jobs or businesses. Sample Adults who refused or did not know how many hours they usually work were 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 were asked the main reason they were not working for pay at a job or business last week. Sample Adults who were unemployed, laid off, looking for work, performing seasonal or contract work, retired, unable to work for health reasons/disabled, taking care of the house or family, going to school, or had some other reason, as well as refused or don't know responses were 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 in the past 12 months, or were working, but not for pay – and all other Sample Adults working within the past 12 months were asked whether paid sick leave is/was available and whether health insurance is/was offered through their workplace , as well as how many days of work they missed because of illness, injury or disability in the past year.

Employment Recodes

Due to changes in the question order and universe in the EMP section starting in 2021, employment recode names have been updated.

Employment status. To ensure that the variables in this section are internally consistent with one another, a summary recode, EMPWRKLSW1_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; those performing seasonal or contract work in the past 12 months; and those doing unpaid work. This recode updates EMPWRKLSW_A from the 2019-2020 NHIS in that it now only includes those who reported seasonal/contract work if they worked in the past 12 months.

Hours Worked. Recode EMPWKHRS3_A provides the hours worked last week top-coded at 95 for confidentiality. Recode EMPWRKFT1_A identifies all employed Sample Adults who usually worked 35 or more hours per week. The two recodes about hours worked update those from 2019-2020 in that these recodes only include seasonal or contract workers if they worked in the past 12 months, and these seasonal/contract workers along with those who worked but not for pay at a family business are no longer coded as not ascertained because they are now asked about their work hours.

Last worked. EMPLSTWOR1_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 unpaid work are coded as "8" or "not ascertained" on this recode. This recode differs from 2019-2020 in that those performing seasonal/contract work are no longer coded as not ascertained because they now provide information on when they last worked.

Workdays missed. To protect confidentiality, information on the number of workdays missed due to health reasons is top-coded at 130 days, EMPDYSMSS3_A. This recode differs from 2019-2020 in that it no longer includes all seasonal/contract workers in the variable universe but only seasonal/contract workers who worked in the past 12 months.

Marital Status

The MAR section contains information on the marital status of the Sample Adult. Sample Adults are first asked if they are "now married, living with a partner together as an unmarried couple, or neither." Sample Adults who responded they are married are asked if their spouse lives in the same residence. If not living in the same residence, they are asked if this is because the Sample Adult and their spouse are legally separated. Sample Adults who are married or living with a partner are asked to verify the sex of their spouse or partner that was obtained during rostering. Sample Adults whose answer to the initial marriage and cohabitation question was other than "married," 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 married nor living with a partner but have been married are asked if they are now widowed, divorced, or separated.

Additional information about the spouse or partner of the Sample Adult is described under "Characteristics about the Spouse or Partner of the Sample Adult."

Parental Status

The MAR section also includes recodes which describe whether the Sample Adult is a parent of a child residing in the family (PARSTAT_A) and the Sample Adult's relationship to Sample Child (SAPARENTSC_A).

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 Veteran Affairs (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.

Rotating Core

Industry and Occupation

Employed Sample Adults – those who were working last week, temporarily absent last week, or were working, but not for pay – and those working within the past 12 months were asked about their occupation, industry, and work activities through a series of questions in the EMD section allowing for verbatim text responses. Sample Adults were first asked for whom they work(ed) at their main job, followed by questions about the kind of business or industry of this job, the kind of work they perform(ed), and their job duties or most important activities. Sample Adults were also asked whether they supervise(d) other employees as part of their job as well as to classify their main job into one of the following categories: employee of a private company, federal government employee, state government employee, local government employee, self-employed, or working without pay at a family-owned business or farm. These detailed employment questions rotate for two consecutive years every three years in the NHIS, starting in 2020. These questions differ from detailed employment questions asked in the 2020 NHIS under the EMP section in that the universe no longer includes seasonal/contract workers unless they worked in the past 12 months.

Industry and Occupation Recodes

The verbatim responses obtained from employed Sample Adults regarding their industry and occupation were coded by statistical clerks and assistants from the Coding & Research Section of the Geography Branch in the National Processing Center at the U.S. Census Bureau. The industry and occupation codes developed by U.S.

Census Bureau for use in non-economic Federal surveys are 4-digit Census codes for industry and occupation consistent with the 2017 North American Industry Classification System (NAICS) and 2018 Standard Occupational Classification (SOC). However, these are not the actual NAICS and SOC codes. Verbatim responses from the Sample Adult are not available to the public.

Available in the Sample Adult public-use files are the detailed occupation recode (EMDOCCUPN1_A) with 94 distinct occupation subgroups, and the associated simple recode (EMDOCCUPN2_A) with 23 major occupation groups. These categories are derived from the 2018 SOC Occupation Subgroups and Major Occupation Groups, respectively, as determined by the U.S. Census Bureau and the Bureau of Labor Statistics. Also available are the detailed industry recode (EMDINDSTN1_A) informed by the 2017 NAICS with 79 distinct industry subsectors, and the associated simple recode (EMDINDSTN2_A) with 21 industry sectors. These are derived from the NAICS Industry Subsectors and Sectors, respectively, as identified by the U.S. Census Bureau. Recode names were updated in 2021 because of the change in universe mentioned above.

The Occupation and Industry Appendices in the Sample Adult Codebook include the response categories and labels to EMDOCCUPN1_A, EMDOCCUPN2_A, EMDINDSTN1_A, and EMDINDSTN2_A. Links to information on the NAICS Industry Subsectors and Sectors and the SOC Occupation Subgroups and Major Groups are embedded within these appendices and provide the classification framework for the recodes on the public-use data file. These lists should not be used in place of the Occupation and Industry Appendices. For more information about the 2017 NAICS, please refer to https://www.bls.gov/ces/naics/home.htm. For more information about the 2018 SOC, please refer to https://www.bls.gov/soc/2018/home.htm.

Sponsored Content

NIOSH sponsored nine questions in the JOB section about shiftwork, schedule flexibility and predictability, income variability or uncertainty, employer-withheld tax deductions, psychosocial job stressors and job insecurity, and presenteeism and absenteeism. The objective of these questions is to describe various work arrangements and schedules as well as psychosocial job stressors to study the links between these factors and health outcomes.

Adults who were privately employed or employed by local, state, or federal government in the past year were asked if taxes are deducted from their pay. All Sample Adults working within the past 12 months were asked about how much their earnings change on a monthly basis, their typical hours of work (daytime, evening, night, rotating, or some other shift), how easy it is to change their work schedule when needed, whether their work schedule changed on a regular basis, and the likelihood of losing their job in the next 12 months. Sample Adults who had regularly changing schedules were also asked how much advanced notice they receive about their scheduled work hours. Sample Adults working within the past 12 months, except for seasonal and contract workers, were asked about the number of days they worked while physically ill in the past 3 months, and those who worked zero to less than 90 days while ill, were also asked about number of missed workdays due to a possible COVID-19 infection.

Characteristics about the Parents of the Sample Child

Annual Core

Starting in 2019, the PAR section includes variables and recodes describing the demographic characteristics of parents residing with the Sample Child, type of relationship (biological, adoptive, step, foster) between the child and his or her parent(s), each parent's current and legal marital status, and information on nativity (whether the parent was born in the U.S. or a U.S. territory).

Detailed information on the type of relationship between Sample Child and parent 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, while the redesigned NHIS allows for the collection of information on up to four parents as long as they all reside in the same household as the child, only information on the first two are released in the public-use file. Instances where a Sample Child had three or four residential parents are rare. Detailed information for the parents residing with the Sample Child is available through the NCHS RDC. For a list of restricted variables, see the PAR section of the Sample Child Codebook for restricted variables.

Several recodes are available in the PAR section of 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). Table 24 lists the available recodes, which include parental sex, age (bottom-coded at 20 and top-coded at 65), education of the parent with the highest educational attainment (top-coded to combine professional school degree and doctoral degree), employment status (including full- versus part-time work), current marital status, legal marital status, and whether the Sample Child's parents are of the same sex or of opposite sex.

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 24. Annual core content of public-use recodes describing demographic characteristics of the parents residing with the Sample Child

Description	Codebook section	Residential parent 1 variable	Residential parent 2 variable	Variable for both residential parents
Relationship type to Sample Child[1]	PAR	RELCHPARENTP1_C	RELCHPARENTP2_C	
Sex	PAR	PARSEX1_C	PARSEX2_C	
Age	PAR	PARAGETC1_C	PARAGETC2_C	
Education of the Sample Child's parent with the highest education	PAR			MAXPAREDUP_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

[1]In 2021, the variables for categorizing the type of relationship between the Sample Child and each residential parent were updated to more clearly identify when there was no residential parent or only one residential parent in the household. These variables of RELCHPARENTP1_C and RELCHPARENTP2_C replace 2019–2020 variables RELCHPARP1_C and RELCHPARP2_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 MAR section includes several recodes describing the 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 Sample Adult module (i.e., MAR, GEN, HHC). These recodes are shown in Table 25 and include the spouse or partner's sex, age (bottom-coded at 20 and top-coded at 85 years), education (top-coded to combine professional school degree and doctoral degree), current work status, and whether the spouse or partner worked full-time (35 or more hours per week).

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 and Household of the Sample Adult and Sample Child."

Table 25. Annual core content of public-use recode variables describing demographic characteristics of the spouse or partner residing with the Sample Adult

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	SPOUSEDUCP_A	PRTNREDUCP_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 26 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 a 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 family members in various age groups in the household.

Table 26. Annual core content of public-use recodes of family and household composition and counter variables available for Sample Adult and Sample Child

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 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 (top-coded)	FAM	MAXEDUCP_A	MAXEDUCP_C
Flag indicating Sample Adults / Sample Child living in households containing more than one family	FAM	MLTFAMFLG_A	MLTFAMFLG_C
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

Family Employment

The family employment section (FEM) contains 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 27.

Table 27. Annual core content of employment counter variables available in the Sample Adult or Sample Child files

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

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 28, are intended to approximate the U.S. Census Bureau's weighted average poverty thresholds for 2020. Because these values were not available when the 2021 NHIS instrument was created, the poverty thresholds used in the 2021 instrument were derived by NCHS from the 2019 poverty thresholds by size of family and number of related children under 18 years, the average Consumer Price Index for all urban consumers (CPI-U) from 2019, the forecasted annual growth rate of the CPI-U for 2020, actual monthly CPI values (all consumers) for January-July 2020, and projected CPI values (all consumers) for August-December 2020.

Table 28. Poverty thresholds used in the 2021 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	$14,000	$19,000	$27,000	$34,000	$55,000
1 person ≥ 66 years	$13,000	$17,000	$25,000	$32,000	$51,000
2 persons, both < 66	$18,000	$24,000	$35,000	$44,000	$71,000
2 persons, 1 is ≥ 66	$16,000	$22,000	$32,000	$40,000	$64,000
3 persons	$21,000	$29,000	$42,000	$52,000	$84,000
4 persons	$27,000	$37,000	$54,000	$67,000	$108,000
5 persons	$32,000	$44,000	$64,000	$80,000	$128,000
6 persons	$36,000	$50,000	$72,000	$91,000	$145,000
7 persons	$41,000	$57,000	$83,000	$103,000	$165,000
8 persons	$46,000	$63,000	$92,000	$115,000	$183,000
9 or more persons	$55,000	$75,000	$109,000	$136,000	$218,000

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 poverty threshold, or if the family's income was greater than or equal to that same poverty threshold.

In 2021, 76.6% of Sample Adults and 82.1% of Sample Child respondents provided their family income. Missing family income for Sample Adults (23.4%) and Sample children (17.9%) were imputed using family income bracket responses or other survey information. Missingness on family income is not completely at random and excluding observations with missing income information can result in biased analyses. Reported and imputed family income information are used to create poverty ratio (see, Recodes of Family Income and Imputed Family Income, below).

To protect confidentiality, continuous and categorical family income, reported and imputed, 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.

Ten multiply imputed income data values are created for the Sample Adult and the Sample Child's families. 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). The family income is imputed within the lower and upper bound when the income bracketing questions are answered.

Ten sets of top-coded continuous poverty ratios and grouped poverty ratios are available for the Sample Adult and Sample Child families. These recodes incorporate information from reported and imputed total family income and are included in the respective income files (Table 29) as a continuous ratio of total family income and family size relative to the poverty threshold and as a finite number of categories.

Table 29. Top-coded variables for family poverty ratio, and related flags available in the Sample Adult and Sample Child files

Description	Sample Adult file (adultinc21)	Sample Child file (childinc21)
Top-coded poverty ratio	POVRATTC_A	POVRATTC_C
Grouped poverty ratio	RATCAT_A	RATCAT_C
Imputed income flag	IMPINCFLG_A	IMPINCFLG_C
Imputation number*	IMPNUM_A	IMPNUM_C

***This variable was named IMPNUM in prior survey years.**

In cases where the Sample Adult and Sample Child are in the same family, these corresponding values are identical. The poverty ratio variable is top-coded to the mean of the top 95th percentile of the 20 distinct imputations from the Sample Adult and Sample Child inhouse poverty ratio values (i.e., 10 poverty ratio values from the Sample Adult and 10 poverty ratio values from the Sample Child).

All 10 imputations are stacked in a single file with a variable (IMPNUM_A or IMPNUM_C) indicating the imputation number in each file. Stacking the 10 multiple imputation datasets into one allows for fewer steps in data preparation for analyses in SAS and Stata. (For sample code on how to use multiply imputation in analysis, see the section 'Merging Survey Data and Imputed Income Files.').

For the convenience of analyses that don't need or use multiple imputed data, the same variables, i.e., POVRATTC_A and RATCAT_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.

> **Analysts are reminded that imputed income files for each year should be merged with the relevant data files for that year before concatenating data files for multiple years.**

For technical information about the imputation model, please refer to the "Imputed Income Technical Document" available with the annual file releases on the NHIS website, under "Using the NHIS.".

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 combine 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.

Telephone Use

The telephone use section (TEL) contains information regarding the availability of a working non-cellular telephone (land line) and of a cellular (wireless, mobile) telephone in the home, and for adults with both types of working telephones, whether the calls received are answered “all or almost all,” “some,” or “very few or none” on the cell phone. The purpose of the telephone questions is to track the percentages of wireless-only adults and of children living in homes with only wireless telephones. The telephone use data are in the Paradata file. Variable information is found in the TEL sections (adult and child) of the Paradata codebook.

Telephone Use Recodes

Recodes that classify the telephone status of the household (HH_PLASS) and the Sample Adult (AD_PLASS) are included in the Paradata file. The response categories in the recodes correspond to the classification presented in the NCHS report "Wireless Substitution: Early Release of Estimates from the National Health Interview Survey" https://www.cdc.gov/nchs/nhis/erwirelesssubs.htm.

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 they 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 secure 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)

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:
    - Medicare (https://www.cdc.gov/nchs/data-linkage/medicare.htm)
    - 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)
- Summer 2022: Department of Veterans Affairs (VA) administrative data, including use of VA benefits (https://www.cdc.gov/nchs/data-linkage/index.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. 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 Beneficiary ID (MBI) for participants aged 65 and older. Additionally, those who refused to provide the last four digits of their SSN or MBI, 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 MBI or provided consent for linkage without SSN or MBI 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, and associated variables are stored in the public use Paradata file.

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., county level data, 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:
    - NCHS-CMS Medicare (https://www.cdc.gov/nchs/data-linkage/medicare-feasibility.htm)
    - NCHS-CMS Medicaid (https://www.cdc.gov/nchs/data-linkage/medicaid-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)
- Geocoded data (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.

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.

American Psychological Association. (2022). Life satisfaction. APA Dictionary of Psychology. Accessed 5/9/2022. Available at https://dictionary.apa.org/life-satisfaction

Bansal R, Gubbi S, Koch CA. (2022). COVID-19 and chronic fatigue syndrome: An endocrine perspective. Journal of Clinical and Translational Endocrinology; 27:100284.

Cheung F., Lucas RE. (2014). Assessing the validity of single-item life satisfaction measures: Results from three large samples. Quality of Life Research, 23(10), 2809-2818.

DuBois DL., Silverthorn N. (2005). Natural Mentoring Relationships and Adolescent Health: Evidence from a National Study, American Journal of Public Health, 95(3), 518-524.

Felitti VJ., Anda RF., Nordenberg D., et al. (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.

Furukawa TA, Kessler RC, Slade T, Andrews G. (2003). The performance of the K6 and K10 screening scales for psychological distress in the Australian National Survey of Mental Health and Well-Being. Psychological medicine, 33(2):357-362.

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

Kessler RC, Andrews G, Colpe LJ, et al. (2002) Short screening scales to monitor population prevalences and trends in non-specific psychological distress. Psychological medicine, 32(6):959-976.

Kessler RC, Barker PR, Colpe LJ, et al. (2003). Screening for serious mental illness in the general population. Arch Gen Psychiatry, 60(2):184–189.

Kessler RC, Green JG, Gruber MJ, et al. (2010). Screening for serious mental illness in the general population with the K6 screening scale: results from the WHO World Mental Health (WMH) survey initiative. International Journal of Methods in Psychiatric Research, 19 Suppl 1(Suppl 1):4-22.

Korn, E.L., Graubard, B.I. (1999). Analysis of Health Surveys, John Wiley & Sons.

Lin JS, Perdue LA, Henrikson NB, Bean SI, Blasi PR. Screening for Colorectal Cancer: An Evidence Update for the U.S. Preventive Services Task Force [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2021 May. Report No.: 20-05271-EF-1.

Melnikow J, Henderson JT, Burda BU, et al. (2018). Screening for Cervical Cancer With High-Risk Human Papillomavirus Testing: A Systematic Evidence Review for the U.S. Preventive Services Task Force [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US), Report No.: 17-05231-EF-1. PMID: 30256575.

National Survey of Children's Health. NSCH Guide to Topics & Questions. Accessed 5/24/2022. Available from https://www.childhealthdata.org/learn-about-the-nsch/topics_questions

Ochiai E, Kigenyi T, Sondik E, et al. (2021). Healthy People 2030 Leading Health Indicators and Overall Health and Well-being Measures: Opportunities to Assess and Improve the Health and Well-being of the Nation. Journal of Public Health Management and Practice, 27(1), S235-S41.

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

Organisation for Economic Co-operation and Development. (2013). OECD Guidelines on Measuring Subjective Well-being. OECD Publishing. http://dx.doi.org/10.1787/9789264191655-en

Schneider SK, O'Donnell L, Stueve A, Coulter RWS. (2012). Cyberbulling, School Bullying, and Psychological Distress: A regional Census of High School Students. American Journal of Public Health, 102(1):171-177.

Sheldrick RC, Henson BS, Neger EN, et al. (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.

Siu AL; U.S. Preventive Services Task Force. Screening for Breast Cancer: U.S. Preventive Services Task Force Recommendation Statement. (2016). Annals of Internal Medicine, 16;164(4):279-96.

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.

Toth, D. (2019). rpms: An R Package for Modeling Survey Data with Regression Trees. Available at: https://mran.microsoft.com/snapshot/2018-08-19/web/packages/rpms/vignettes/rpms_2018_01_22.pdf

US Preventive Services Task Force. (2022). Aspirin Use to Prevent Cardiovascular Disease: US Preventive Services Task Force Recommendation Statement. JAMA, 26;327(16):1577-1584.

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): 2021 NHIS

Module	Section Acronym	Questionnaire Variable	Description	Availability	Public-use Recode
Roster	HHC	EDUC	Highest level of education completed	D R	EDUCP_A MAXPAREDUP_C SPOUSEDUCP_A PRTNREDUCP_A MAXEDUCP_A MAXEDUCP_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, BMICAT_A
Sample Adult	BMI	WEIGHTLB_A, WEIGHTKG_A	Weight	D R	WEIGHTLBTC_A, BMICAT_A
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 THROACAN_A THYROCAN_A UTERUCAN_A HDNCKCAN_A COLRCCAN_A OTHERCANP_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 Child	CVV	CVDVAC1M_C	Month of most recent COVID-19 vaccination	D R	CVDVAC1MR_C
Sample Child	CVV	CVDVAC1Y_C	Year of most recent COVID-19 vaccination	D R	CVDVAC1YR_C
Sample Child	CVV	CVDVAC2M_C	Month of next most recent COVID-19 vaccination	D R	CVDVAC2MR_C
Sample Child	CVV	CVDVAC2Y_C	Year of most recent COVID-19 vaccination	D R	CVDVAC2YR_C
Sample Adult	DIB	DIBAGE_A	Age of diabetes diagnosis	D R	DIBAGETC_A, DIFYRSTC1_A
Sample Adult	EMP	EMPLASTWK_A	Worked for pay last week	D R	EMPWRKLSW1_A
Sample Adult	EMP	EMPNOWRK_A	Temporarily absent from work last week	D R	EMPWRKLSW1_A
Sample Adult	EMP	EMPHOURS_A	Hours worked per week	D R	EMPWKHRS3_A, EMPWRKFT1_A
Sample Adult	EMP	EMPFULLTIM_A	Work 35+ hours per week	D R	EMPWRKFT1_A
Sample Adult	EMP	EMPWHENWRK_A	Last time worked	D R	EMPLSTWOR1_A
Sample Adult	EMP	EMPDAYMISS_A	Days missed work, past 12 months	D R	EMPDYSMSS3_A
Sample Adult	EMD	EMDWHOWRK1_A	Name of company, business or employer	NA	
Sample Adult	EMD	EMDINDUST_A	Kind of business or industry	R	EMDINDSTRN_A, EMDINDSTN1_A, EMDINDSTN2_A
Sample Adult	EMD	EMDKIND1_A	Kind of work doing	R	EMDOCCUPN_A, EMDOCCUPN1_A, EMDOCCUPN2_A
Sample Adult	EMD	EMDIMPACT1_A	Most important job activities or duties	NA	
Sample Adult	EPI	EPINUMSEZ_A	Number of seizures	D R	EPINUMSEZP_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	CVDVAC1M_A	Month of most recent COVID-19 vaccination	D R	CVDVAC1MR_A
Sample Adult	IMS	CVDVAC1Y_A	Year of most recent COVID-19 vaccination	D R	CVDVAC1YR_A
Sample Adult	IMS	CVDVAC2M_A	Month of next most recent COVID-19 vaccination	D R	CVDVAC2MR_A
Sample Adult	IMS	CVDVAC2Y_A	Year of next most recent COVID-19 vaccination	D R	CVDVAC2YR_A
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	INC	INCTOTAL_A	Total family income as poverty ratio	D R	POVRATTC_A, RATCAT_A
Sample Adult	INC	INCTOTAL_A INC100FILL_A INC250FILL_A INC138FILL_A INC75K_A INC100K_A	Reported and imputed Sample Adult family income	D	

		INC138K_A INC150K_A INC250K_A INC400K_A INC100PCT_A INC138PCT_A INC200PCT_A INC250PCT_A INC400PCT_A			
Sample Child	INC	INCTOTAL_C	Total family income as poverty ratio	D R	POVRATTC_C, RATCAT_C
Sample Child	INC	INCTOTAL_C PCNT_C INC100FILL_C INC250FILL_C INC138FILL_C INC75K_C INC100K_C INC138K_C INC150K_C INC250K_C INC400K_C INC100PCT_C INC138PCT_C INC200PCT_C INC250PCT_C INC400PCT_C	Reported and imputed Sample Child family income	D	
Sample Adult	INJ	NUMINJ_A	Number of non-repetitive strain injuries in the past 3 months	D R	NUMINJTC_A
Sample Adult	INJ	INJWRKDAYS_A	Number of workdays missed due to injury in the past 3 months	D R	INJWRKDYTC_A
Sample Child	INJ	INJSCHDAYS_C	Number of school days missed due to injury in the past 3 months	D R	INJSCHDYTC_C
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	RSNHIOTHSP_A	Verbatim response to reasons for not getting coverage	R	RSNHICOST_A, RSNHIWANT_A, RSNHIELIG_A, RSNHICONF_A, RSNHIMEET_A, RSNHIWAIT_A, RSNHIOTH_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, HIKIND01R_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	POLHLDA1_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	POLHLDA2_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	LS1	LSATIS11_A	Life satisfaction-number	D R	LSATIS11R_A
Sample Adult	LS3	LSATIS4_A	Life satisfaction/dissatisfaction	D R	LSATIS4R_A
Sample Adult	MAR	SPOUSWHO_A	Person number for spouse	D	
Sample Adult	MAR	SPOUSSEX_A, SPOUNEWSEX_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_	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	REP	REPWRKDAYS_A	Numbers workdays missed due to repetitive strain injury	D R	REPWRKDYTC_A
Sample Adult	PRV	DIBA1CNUM_A	Number of times A1C checked	D R	DIBA1CNMT_A
Sample Adult	PRV	COLTEST_A	Colon test recommended	D R	COLTEST1_A- COLTEST6_A
Sample Adult	SCH	SCHDYSMSS_A	Number of school days missed, past 12m	D R	SCHDYMSSTC_A
Sample Child	SCH	SCHDYSMSS_C	Number of school days missed, past 12m	D R	SCHDYMSSTC_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/.