

**MEPS HC-233
2021 Full Year
Consolidated Data File**

August 2023

**Agency for Healthcare Research and Quality
Center for Financing, Access, and Cost Trends
5600 Fishers Lane
Rockville, MD 20857
(301) 427-1406**

Table of Contents

<u>Section</u>		<u>Page</u>
A.	Data Use Agreement.....	A-1
B.	Background.....	B-1
	1.0 Household Component.....	B-1
	2.0 Medical Provider Component.....	B-1
	3.0 Survey Management and Data Collection	B-2
C.	Technical and Programming Information.....	C-1
	1.0 General Information.....	C-1
	2.0 Data File Information.....	C-2
	2.1 Codebook Structure	C-4
	2.2 Reserved Codes	C-5
	2.3 Codebook Format	C-5
	2.4 Variable Naming	C-6
	2.5 File Contents.....	C-6
	2.5.1 Survey Administration Variables (DUID - RURSLT53)	C-7
	2.5.2 Navigating the MEPS Data with Information on Person Disposition Status	C-17
	2.5.3 Demographic Variables (AGE31X-DAPID53X)	C-22
	2.5.4 Income and Tax Filing Variables (FILEDR21-HIEUIDX).....	C-33
	2.5.5 Person-Level Condition Variables (RTHLTH31-ADHDAGED)	C-38
	2.5.6 Health Status Variables (IADLHP31-ADOVER42)	C-43
	2.5.7 Disability Days Indicator Variables (DDNWRK21-OTHNDD21).....	C-66
	2.5.8 Access to Care Variables (ACCELI42- BOOSTERSHOT53).....	C-67
	2.5.9 Employment Variables (EMPST31-RTPLN53H)	C-73
	2.5.10 Health Insurance Variables (TRIJAyyX-PMEDPY53).....	C-95

<u>Section</u>	<u>Page</u>
2.5.11 Utilization, Expenditures, and Sources of Payment Variables (TOTTC21-RXOSR21).....	C-112
2.5.12 Changes in Variable List.....	C-121
2.6 Linking to Other Files	C-125
2.6.1 Event and Condition Files.....	C-125
2.6.2 National Health Interview Survey.....	C-125
2.6.3 Longitudinal Analysis	C-125
3.0 Survey Sample Information	C-126
3.1 Discussion of Pandemic Effects on Quality of 2021 MEPS Data	C-126
3.1.1 Summary	C-126
3.1.2 Overview	C-126
3.1.3 Data Quality Issues for MEPS in 2021 Directly Associated with Data Quality Concerns for the NHIS and CPS.....	C-126
3.1.4 Modifications to the MEPS HC 2021 Sample Design	C-127
3.1.5 Data Quality Issues for MEPS for FY 2021.....	C-128
3.1.6 Discussion and Guidance	C-129
3.2 Background on Sample Design and Response Rates	C-130
3.2.1 MEPS-Linked to the National Health Interview Survey (NHIS).....	C-131
3.2.2 Sample Weights and Variance Estimation	C-133
3.3 The MEPS Sampling Process and Response Rates: An Overview	C-133
3.3.1 Response Rates	C-135
3.3.2 Panel 26 Response Rates.....	C-136
3.3.3 Panel 25 Response Rates.....	C-136

<u>Section</u>	<u>Page</u>
3.3.4	Panel 24 Response Rates..... C-137
3.3.5	Panel 23 Response Rates..... C-137
3.3.6	Annual (Combined Panel) Response Rate C-137
3.3.7	Oversampling C-137
3.4	Background on Person-Level Estimation Using this MEPS Public Use Release C-139
3.4.1	Requirements to Receive a Person- Level Weight..... C-139
3.4.2	Details on Person-Level Weights Construction C-139
3.4.3	MEPS Panel 23 Weight Development Process C-140
3.4.4	MEPS Panel 24 Weight Development Process C-141
3.4.5	MEPS Panel 25 Weight Development Process C-141
3.4.6	MEPS Panel 26 Weight Development Process C-142
3.4.7	The Final Person-Level Weight for 2021 C-142
3.4.8	A Note on MEPS Population Estimates C-143
3.5	Coverage..... C-143
3.6	Background on Family-Level Estimation Using This MEPS Public Use File C-144
3.6.1	Overview C-144
3.6.2	Definition of “Family” for Estimation Purposes C-144
3.6.3	Instructions to Create Family Estimates C-145
3.6.4	Details on Family Weight Construction and Estimated Number of Families..... C-146
3.7	Analysis Using Health Insurance Eligibility Units C-147

<u>Section</u>	<u>Page</u>	
3.8	Weights and Response Rates for the Self-Administered Questionnaire	C-148
3.9	Weights and Response Rates for the Diabetes Care Survey	C-149
3.10	Weights and Response Rates for the Social Determinants of Health Survey	C-150
3.11	Variance Estimation	C-151
3.11.1	Taylor-series Linearization Method.....	C-151
3.11.2	Balanced Repeated Replication (BRR) Method.....	C-153
3.12	Guidelines for Determining which Weight to Use for Analyses Involving Data/Variables from Multiple Sources and Supplements: MEPS 2021 Full-Year Use File.....	C-153
3.13	Using MEPS Data for Trend Analysis	C-154
	References.....	C-156
D.	Variable-Source Crosswalk	D-1
 <u>Appendices</u>		<u>Page</u>
1	MEPS Industry Codes Condensing Rules	A1-1
2	MEPS Occupation Codes Condensing Rules	A2-1
3	Summary of Utilization and Expenditure Variables by Health Service Category.....	A3-1

A. Data Use Agreement

Individual identifiers have been removed from the micro-data contained in these files. Nevertheless, under sections 308 (d) and 903 (c) of the Public Health Service Act (42 U.S.C. 242m and 42 U.S.C. 299 a-1), data collected by the Agency for Healthcare Research and Quality (AHRQ) and/or the National Center for Health Statistics (NCHS) may not be used for any purpose other than for the purpose for which they were supplied; any effort to determine the identity of any reported cases is prohibited by law.

Therefore in accordance with the above referenced Federal Statute, it is understood that:

1. No one is to use the data in this data set in any way except for statistical reporting and analysis; and
2. If the identity of any person or establishment should be discovered inadvertently, then (a) no use will be made of this knowledge, (b) the Director Office of Management AHRQ will be advised of this incident, (c) the information that would identify any individual or establishment will be safeguarded or destroyed, as requested by AHRQ, and (d) no one else will be informed of the discovered identity; and
3. No one will attempt to link this data set with individually identifiable records from any data sets other than the Medical Expenditure Panel Survey or the National Health Interview Survey. Furthermore, linkage of the Medical Expenditure Panel Survey and the National Health Interview Survey may not occur outside the AHRQ Data Center, NCHS Research Data Center (RDC) or the U.S. Census RDC network.

By using these data you signify your agreement to comply with the above stated statutorily based requirements with the knowledge that deliberately making a false statement in any matter within the jurisdiction of any department or agency of the Federal Government violates Title 18 part 1 Chapter 47 Section 1001 and is punishable by a fine of up to \$10,000 or up to 5 years in prison.

The Agency for Healthcare Research and Quality requests that users cite AHRQ and the Medical Expenditure Panel Survey as the data source in any publications or research based upon these data.

B. Background

1.0 Household Component

The Medical Expenditure Panel Survey (MEPS) provides nationally representative estimates of health care use, expenditures, sources of payment, and health insurance coverage for the U.S. civilian noninstitutionalized population. The MEPS Household Component (HC) also provides estimates of respondents' health status, demographic and socio-economic characteristics, employment, access to care, and satisfaction with health care. Estimates can be produced for individuals, families, and selected population subgroups. The panel design of the survey includes five rounds of interviews covering two full calendar years. Additional rounds were added in 2020 and 2021, covering third and fourth years respectively, to compensate for the smaller number of completed interviews in later panels. These extra rounds provide data for examining person-level changes in selected variables such as expenditures, health insurance coverage, and health status. Using computer assisted personal interviewing (CAPI) technology, information about each household member is collected, and the survey builds on this information from interview to interview. All data for a sampled household are reported by a single household respondent.

The MEPS HC was initiated in 1996. Each year a new panel of sample households is selected. Because the data collected are comparable to those from earlier medical expenditure surveys conducted in 1977 and 1987, it is possible to analyze long-term trends. Each annual MEPS HC sample size is about 15,000 households. Data can be analyzed at either the person or event level. Data must be weighted to produce national estimates.

The set of households selected for each panel of the MEPS HC is a subsample of households participating in the previous year's National Health Interview Survey (NHIS) conducted by the National Center for Health Statistics. The NHIS sampling frame provides a nationally representative sample of the U.S. civilian noninstitutionalized population. In 2006, the NHIS implemented a new sample design, which included Asian persons in addition to households with Black and Hispanic persons in the oversampling of minority populations. NHIS introduced a new sample design in 2016 that discontinued oversampling of these minority groups.

2.0 Medical Provider Component

Upon completion of the household CAPI interview and obtaining permission from the household survey respondents, a sample of medical providers are contacted by telephone to obtain information that household respondents can not accurately provide. This part of the MEPS is called the Medical Provider Component (MPC) and information is collected on dates of visits, diagnosis and procedure codes, charges and payments. The Pharmacy Component (PC), a subcomponent of the MPC, does not collect charges or diagnosis and procedure codes but does collect drug detail information, including National Drug Code (NDC) and medicine name, as well as amounts of payment. The MPC is not designed to yield national estimates. It is primarily used as an imputation source to supplement/replace household reported expenditure information.

3.0 Survey Management and Data Collection

MEPS HC and MPC data are collected under the authority of the Public Health Service Act. Data are collected under contract with Westat, Inc. (MEPS HC) and Research Triangle Institute (MEPS MPC). Data sets and summary statistics are edited and published in accordance with the confidentiality provisions of the Public Health Service Act and the Privacy Act. The National Center for Health Statistics (NCHS) provides consultation and technical assistance.

As soon as data collection and editing are completed, the MEPS survey data are released to the public in staged releases of micro data files and tables via the [MEPS website](#) and datatools.ahrq.gov.

Additional information on MEPS is available from the MEPS project manager or the MEPS public use data manager at the Center for Financing, Access, and Cost Trends, Agency for Healthcare Research and Quality, 5600 Fishers Lane, Rockville, MD 20857 (301-427-1406).

C. Technical and Programming Information

1.0 General Information

This documentation describes the 2021 Full-Year Consolidated Data File from the Medical Expenditure Panel Survey Household Component (MEPS HC). Released as an ASCII file (with related SAS, Stata, SPSS, and R programming statements and data user information), a SAS data set, a SAS transport dataset, a Stata data set, and an Excel file this public use file provides information collected on a nationally representative sample of the civilian noninstitutionalized population of the United States for calendar year 2021. The file contains 1,488 variables and has a logical record length of 4,287 with an additional 2-byte carriage return/line feed at the end of each record.

This file consists of MEPS survey data obtained in the 2021 portion of Round 7, and all of Rounds 8 and 9 for Panel 23; the 2021 portion of Rounds 5, and 7, and all of Round 6 for Panel 24; the 2021 portion of Round 3, and all of Rounds 4 and 5 for Panel 25; and Rounds 1, 2, and the 2021 portion of Round 3 for Panel 26 (i.e., the rounds for the MEPS panels covering calendar year 2021). 2021 is the first data year to include four panels of data; Panel 23 was extended to include Rounds 7, 8, and 9 and Panel 24 was extended to include Rounds 6 and 7. In addition, the Panel 24 Round 5 reference period was extended into 2021 instead of ending on 12/31/2020. An overview of the impact of this change on variables and variable names is described in Section 2.4.

This file contains variables pertaining to survey administration, demographics, income, person-level conditions, health status, disability days, quality of care, employment, health insurance, and person-level medical care use and expenditures.

The following documentation offers a brief overview of the types and levels of data provided, content and structure of the files, and programming information. It contains the following sections:

- Data File Information
- Survey Sample Information
- Variable-Source Crosswalk

Both weighted and unweighted frequencies of most variables included in the 2021 full-year consolidated data file are provided in the accompanying codebook file. The exceptions to this are weight variables, variance estimation variables, and variables that have a separate weight (SAQ, DCS, and SDOH). Only unweighted frequencies of these variables are included in the accompanying codebook file. See the Weights Variables list in Section D, Variable-Source Crosswalk.

A database of all MEPS products released to date can be found on the [MEPS website](#).

2.0 Data File Information

This public use dataset contains variables and frequency distributions associated with 28,336 persons who participated in the MEPS Household Component of the Medical Expenditure Panel Survey in 2021. These persons received a positive person-level weight, a family-level weight, or both (some participating persons belonged to families characterized as family-level nonrespondents while some members of participating families were not eligible for a person-level weight).

These 28,336 persons were part of one of the four MEPS panels for whom data were collected in 2021: Rounds 7, 8, and 9 of Panel 23; Rounds 5, 6, and 7 of Panel 24; Rounds 3, 4, and 5 of Panel 25; or Rounds 1, 2, and 3 of Panel 26. Of these persons, 27,332 were assigned a positive person-level weight. There were 13,238 families receiving a positive family-level weight. The codebook provides both weighted and unweighted frequencies for most variables on the dataset. In conjunction with the person-level weight variable (PERWT21F) provided on this file, data for persons with a positive person-level weight can be used to make estimates for the civilian noninstitutionalized U.S. population for 2021.

The MEPS CAPI design has changed significantly beginning with the specifications for Panel 21 Round 5/Panel 22 Round 3/Panel 23 Round 1. In addition, in 2021, three rounds of data collection were added for Panel 23 (Rounds 7, 8, and 9), Round 5 was extended for Panel 24 and two rounds of data collection were added for Panel 24 (Rounds 6 and 7).

	2019				2020				2021				2022			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Panel 23																
Round 3	[Cyan bar]															
Round 4	[Cyan bar]															
Round 5					[Cyan bar]											
Round 6					[Yellow bar]											
Round 7									[Yellow bar]							
Round 8									[Red bar]							
Round 9													[Red bar]			
Panel 24																
Round 1	[Cyan bar]															
Round 2	[Cyan bar]															
Round 3					[Cyan bar]				[Yellow bar]							
Round 4					[Yellow bar]											
Round 5									[Yellow bar]							
Round 6									[Red bar]							
Round 7									[Red bar]				[Green bar]			
Round 8													[Green bar]			
Round 9													[Green bar]			
Panel 25																
Round 1					[Yellow bar]											
Round 2					[Yellow bar]											
Round 3									[Yellow bar]				[Red bar]			
Round 4									[Red bar]							
Round 5													[Red bar]			
Panel 26																
Round 1									[Red bar]							
Round 2									[Red bar]							
Round 3									[Red bar]				[Green bar]			
Round 4													[Green bar]			
Round 5													[Green bar]			
Panel 27																
Round 1													[Green bar]			
Round 2													[Green bar]			
Round 3													[Green bar]			
Sample Size	N = 27,648				N = 26,847				N = 27,332				N = TBD			

N is equal to the number of people with a positive person weight on the file

The following sections of the CAPI instrument were asked in the new rounds:

Variable Grouping	CAPI Section	Asked in Round 6	Asked in Round 7	Asked in Round 8	Asked in Round 9
Survey Administration and Demographics	ST, RE-A, RE-B, CL	X	X	X	X
Person-Level Conditions	PE	X	X	X	X

Variable Grouping	CAPI Section	Asked in Round 6	Asked in Round 7	Asked in Round 8	Asked in Round 9
Health Status	AH, DCS		X		X
Health Status	CS, SAQ	X		X	
Health Status	HE	X	X	X	
Disability Days	AH	X	X	X	X
Access to Care	AC	X		X	
Access to Care	CV	X	X	X	X
Employment	RJ, EM, EW	X	X	X	X
Health Insurance	HX, OE, HP, HQ, MC, PR	X	X	X	X
Use and Expenditures	CA, EV, PP, ED, EF, DV, ER, HS, MV, OP, HH, PM, TH	X	X	X	X

2.1 Codebook Structure

The codebook and data file sequence lists variables in the following order:

- Unique person identifiers and survey administration variables
- Geographic variables
- Demographic variables
- Income and tax filing variables
- Person-level priority condition variables
- Health status variables
- Disability days variables
- Access to care and COVID variables
- Employment variables
- Health insurance variables
- Utilization, expenditure, and source of payment variables
- Weight and variance estimation variables

2.2 Reserved Codes

The following reserved code values are used:

Value	Definition
-1 INAPPLICABLE	Question was not asked due to skip pattern
-2 DETERMINED IN PREVIOUS ROUND	Question was not asked in round because there was no change in current main job since previous round
-7 REFUSED	Question was asked and respondent refused to answer question
-8 DK	Question was asked and respondent did not know answer or the information could not be ascertained
-10 HOURLY WAGE >= \$105.77	Hourly wage was top-coded for confidentiality
-13 INITIAL WAGE IMPUTED	Hourly wage was previously imputed so an updated wage is not included in this file
-15 CANNOT BE COMPUTED	Value cannot be derived from data

The value -15 (CANNOT BE COMPUTED) is assigned to MEPS constructed variables in cases where there is not enough information from the MEPS instrument to calculate the constructed variables. “Not enough information” is often the result of skip patterns in the data or from missing information resulting from MEPS responses of -7 (REFUSED) or -8 (DK). Note that reserved code -8 includes cases where the information from the question was “not ascertained” or where the respondent chose “don’t know”.

2.3 Codebook Format

This codebook describes an ASCII data set (although the data are also being provided in a SAS data set, SAS transport file, Stata data set, and Excel file) and provides the following programming identifiers for each variable:

Identifier	Description
Name	Variable name
Description	Variable descriptor
Format	Number of bytes
Type	Type of data: numeric (indicated by NUM) or character (indicated by CHAR)
Start	Beginning column position of variable in record
End	Ending column position of variable in record

2.4 Variable Naming

In general, variable names reflect the content of the variable. Edited variables end in an X and are so noted in the variable label. The last two characters in round-specific variables have denoted the rounds of data collection, for example, Round 3, 4, or 5 of Panel 25, and Round 1, 2, or 3 of Panel 26. Historically round dates have been indicated by two numbers following the variable name; the first number representing the round for second panel persons (Panel 25), the second number representing the round for first panel persons (Panel 26). For example, EMPST31 refers to employment status on the Round 3 interview date for Panel 25 persons and employment status on the Round 1 interview date for Panel 26 persons.

The variable names in the 2021 files have not been renamed from prior years, despite the addition of Rounds 5, 6, and 7 of Panel 24 and Rounds 7, 8, and 9 of Panel 23, and those round numbers (5, 6, and 7 for Panel 24 and 7, 8, and 9 for Panel 23) will not be included in the variable names. The variable names containing “53” (for example, AGE53X, INS53X, INSAT53X, EMPST53, and HELD53X) have data from Round 9 of Panel 23, Round 7 of Panel 24, Round 5 of Panel 25, and Round 3 of Panel 26. The variable names containing “42” (for example, AGE42X, INS42X, INSAT42X, EMPST42, and HELD42X) have data from Round 8 of Panel 23, Round 6 of Panel 24, Round 4 of Panel 25, and Round 2 of Panel 26. The variable names containing “31” have data from Round 7 of Panel 23, Round 5 of Panel 24, Round 3 of Panel 25, and Round 1 of Panel 26.

As variable collection, universe, or categories are altered, the variable name will be appended with “_Myy” to indicate in which year the alterations took place. Details about these alterations can be found throughout this document.

Variables contained in this delivery were derived either from the questionnaire itself or from the CAPI. The source of each variable is identified in the section of the documentation entitled “Section D. Variable-Source Crosswalk.” Sources for each variable are indicated in one of four ways: (1) variables derived from CAPI or assigned in sampling are so indicated; (2) variables derived from complex algorithms associated with re-enumeration are labeled “RE Section”; (3) variables that are collected by one or more specific questions in the instrument have those question numbers listed in the Source column; and (4) variables constructed from multiple questions using complex algorithms are labeled “Constructed.”

2.5 File Contents

Users of MEPS data should be aware that the survey collects data for all sample persons who were in the survey target population at any time during the survey period. In other words, a small proportion of individuals in MEPS analytic files are not members of the survey target population (i.e., civilian noninstitutionalized) for the entire survey period. These persons include those who had periods during which they lived in an institution (e.g., nursing home or prison), were in the military, or lived out of the country, as well as those who were born (or adopted) into MEPS sample households or died during the year. They are considered sample persons for the survey and are included in MEPS data files with positive person weights, but no data were collected for the periods they were not in scope and their annual data for variables like health care utilization,

expenditures, and insurance coverage reflect only the part of the year they were inscope for the survey. Persons who are inscope for only part of the year should not be confused with non-respondents. Sample persons who are classified as non-respondents to one or more rounds of data collection (i.e., initial non-respondents and dropouts over time) are not included in MEPS annual files, and survey weights for full-year respondents are inflated through statistical adjustment procedures to compensate for both full and part-year nonresponse (see Section 3.0 “Survey Sample Information” for more information). The AHRQ website provides more details about the [identification and analytic considerations regarding sample persons who are inscope only part of the year](#).

2.5.1 Survey Administration Variables (DUID - RURSLT53)

The survey administration variables contain information related to conducting the interview, household and family composition, and person-level and RU-level status codes. Data for the survey administration variables were derived from the sampling process, the CAPI programs, or were computed based on information provided by the respondent in the Reenumeration section of the questionnaire. Most survey administration variables on this file are asked during every round of the MEPS interview. They describe data for Rounds 7/5/3/1, 8/6/4/2, 9/7/5/3 status and status as of December 31, 2021.

This year, the data collected includes a fourth panel, Panel 23, which was extended past the previous seven rounds of data collection to include eighth and ninth round interviews. In order to incorporate the fourth year Panel 23 Round 7, Round 8, and Round 9 data into the delivery without adding new variables, a decision was made to use the ‘31’/‘42’/‘53’ variables to also hold the fourth year panel data. Retaining the standard ‘31’/‘42’/‘53’ variable names, even with the addition of the fourth year panel, will allow easier comparison to previous FY data. Data collection for Panel 24 was planned to go beyond the usual five rounds, so for 2021, Round 5 was collected as a (2020-2021) cross-year round. Alternatively, Panel 25 data collection was planned to end at five rounds, and thus, Panel 25 Round 5 was collected as a (2021) terminal round (not as a 2021-2022 cross-year round).

For example, where variables ending in ‘53’ would normally hold data for Panel 25 Round 5 and Panel 26 Round 3, they also now hold data for Panel 24 Round 7 and Panel 23 Round 9. Similarly, Panel 24 Round 6 and Panel 23 Round 8 data have been added to the ‘42’ variables. Panel 24 Round 5 and Panel 23 Round 7 data have been added to the ‘31’ variables. This means that the ‘31’ variables contain data for Rounds 1, 3, 5, and 7, the ‘42’ variables contain data for Rounds 2, 4, 6, and 8, and the ‘53’ variables contain data for Rounds 3, 5, 7, and 9.

The December 31, 2021 variables were developed in two ways. Those used in the construction of eligibility, inscope, and the end reference date were based on an exact date. The remaining variables were constructed using data from specific rounds, if available. If data were missing from the target round but were available in another round, data from that other round were used in the variable construction. If no valid data were available during any round of data collection, an appropriate reserved code was assigned.

Dwelling Units, Reporting Units, and Families

The definitions of Dwelling Units (DUs) in the MEPS Household Survey are generally consistent with the definitions employed for the National Health Interview Survey (NHIS). The Dwelling Unit ID (DUID) is a seven-digit ID number consisting of a 2-digit panel number followed by a five-digit random number assigned after the case was sampled for MEPS. A three-digit person number (PID) uniquely identifies each person within the DU. The variable DUPERSID is the combination of the variables DUID and PID. As part of the new CAPI design, the lengths of the ID variables have changed in the file. An additional 2 bytes in the IDs resulted from adding a 2-digit panel number to the beginning of all the IDs. Users should be mindful of the different ID structures/lengths when combining MEPS files from 1996-2017 with MEPS files from 2018-2021.

PANEL is a constructed variable used to specify the panel number for the person. PANEL will indicate Panel 23, Panel 24, Panel 25, or Panel 26 for each person on the file. Panel 23 is the panel that started in 2018, Panel 24 is the panel that started in 2019, Panel 25 is the panel that started in 2020, and Panel 26 is the panel that started in 2021. The panel number is included as the first two digits of the DUID and DUPERSID.

A Reporting Unit (RU) is a person or group of persons in the sampled DU who are related by blood, marriage, adoption, or other family association. Each RU was interviewed as a single entity for MEPS. Thus, the RU serves chiefly as a family-based “survey” operations unit rather than an analytic unit. Members of each RU within the DU are identified in the pertinent three rounds by the round-specific variables RULETR31, RULETR42, and RULETR53. End-of-year status (as of December 31, 2021 or the last round they were in the survey) is indicated by the RULETR21 variable. Regardless of the legal status of their association, two persons living together as a “family” unit were treated as a single RU if they chose to be so identified.

Examples of different types of RUs are:

1. A married daughter and her husband living with her parents in the same DU constitute a single RU;
2. A husband and wife and their unmarried daughter, age 18, who is living away from home while at college constitute two RUs; and
3. Three unrelated persons living in the same DU would each constitute a distinct RU (a total of three RUs).

Unmarried college students (less than 24 years of age) who usually live in the sampled household but were living away from home and going to school at the time of the Round 3/1 MEPS interview were treated as an RU separate from that of their parents for the purpose of data collection.

The round-specific variables RUSIZE31, RUSIZE42, RUSIZE53, and the end-of-year status variable RUSIZE21 indicate the number of persons in each RU, treating students as single RUs separate from their parents. Thus, students are not included in the RUSIZE count of their parents' RU. However, for many analytic objectives, the student RUs would be combined with

their parents' RU, treating the combined entity as a single family. Family identifier and size variables are described below and include students with their parents' RU.

The round-specific variables FAMID31, FAMID42, FAMID53, and the end-of-year status variable FAMID21 identify a family (i.e., persons related to one another by blood, marriage, adoption, or self-identified as a single unit) for each round and as of December 31, 2021. The FAMID variables differ from the RULETR variables only in that student RUs are combined with their parents' RU.

Two other family identifiers, FAMIDYR and CPSFAMID, are provided on this file. The annualized family ID letter, FAMIDYR, identifies eligible members of the eligible annualized families within a DU. The CPSFAMID represents a redefinition of MEPS families into families defined by the Current Population Survey (CPS). Some of the distinctions between CPS-and MEPS-defined families are that MEPS families include and CPS families do not include: non-married partners, and in-laws. These persons are considered as members of separate families for CPS-like families. CPS-like families are defined so a poverty status classification variable consistent with established definitions of poverty can be assigned to the CPS-like families and used for weight poststratification purposes. In order to identify a person's family affiliation, users must create a unique set of FAMID variables by concatenating the DU identifier and the FAMID variable. Instructions for creating family estimates are described in Section 3.6.

Foster care relationships and fostered members of households are no longer included in the MEPS data. This change was implemented with the 2017 Consolidated Public Use File, so users combining many years of data may encounter foster relationships/members in earlier MEPS files.

The round-specific variables FAMSZE31, FAMSZE42, FAMSZE53, and the end-of-year status variable FAMSZE21 indicate the number of persons associated with a single family unit after students are linked to their associated parent RUs for analytical purposes. Family-level analyses should use the FAMSZE variables.

Note that the variables RUSIZE31, RUSIZE42, RUSIZE53, RUSIZE21, FAMSZE31, FAMSZE42, FAMSZE53, and FAMSZE21 exclude persons who are ineligible for data collection (i.e., those where ELGRND31 NE 1, ELGRND42 NE 1, ELGRND53 NE 1 or ELGRND21 NE 1); analysts should exclude ineligible persons in a given round from all family-level analyses for that round.

The round-specific variables RURSLT31, RURSLT42, and RURSLT53 indicate the RU response status for each round. Users should note that the values for RURSLT31 differ from those for RURSLT42 and RURSLT53. The values for RURSLT31 include the following:

Value	Definition
-1	Inapplicable
60	Complete with RU member
61	Complete with proxy-all RU members deceased

Value	Definition
62	Complete with proxy-all RU members institutionalized or deceased
63	Complete with proxy-other
72	RU institutionalized in prior round; Still institutionalized-R3 only
80	Entire RU merged with other RU
81	Entire RU deceased before 1/1/21
82	Entire RU is military before 1/1/21
83	Entire RU institutionalized before 1/1/21
84	Entire RU left U.S. before 1/1/21
85	Entire RU is ineligible before 1/1/21; Multi-reason
86	Entire RU is ineligible; Non-Key NHIS study
87	Reenumeration complete; No eligible RU member; Ineligible RU
88	Unavailable during field period
89	Too ill; No proxy
90	Physically/Mentally incompetent; No proxy
91	Final Refusal
92	Final Breakoff
93	Unable to locate
94	Entire RU is military or left U.S. after 1/1/21
95	Entire RU institutionalized after 1/1/21; No proxy
96	Entire RU deceased after 1/1/21; No proxy
97	Reenumeration complete; No RU member; Non-Response
98	RU moved too far to interview
99	Final other Non-Response

The values for RURSLT42 and RURSLT53 include the following:

Value	Definition
-1	Inapplicable
60	Complete with RU member
61	Complete with proxy-all RU members deceased
62	Complete with proxy-all RU members institutionalized or deceased
63	Complete with proxy-other
70	Entire RU merged with other RU

Value	Definition
71	Reenumeration complete; No eligible RU member; Ineligible RU
72	RU institutionalized in prior round; Still institutionalized
88	Unavailable during field period
89	Too ill; No proxy
90	Physically/Mentally incompetent; No proxy
91	Final Refusal
92	Final Breakoff
93	Unable to locate
94	Entire RU is military or left U.S. after 1/1/21
95	Entire RU institutionalized after 1/1/21; No proxy
96	Entire RU deceased after 1/1/21; No proxy
97	Reenumeration complete; No RU member; Non-Response
98	RU moved too far to interview
99	Final other Non-Response

Standard or primary RUs are the original RUs from NHIS. A new RU is one created when members of the household leave the primary RU and are followed according to the rules of the survey. A student RU is an unmarried college student (under 24 years of age) who is considered a usual member of the household, but was living away from home while going to school, and was treated as a Reporting Unit (RU) separate from their parents' RU for the purpose of data collection. RUCLAS21 was set based on the RUCLAS values from Rounds 7/5/3/1, 8/6/4/2, and 9/7/5/3. If the person was present in the responding RU in Round 9/7/5/3, then RUCLAS21 was set to RUCLAS53. If the person was not present in a responding RU in Round 9/7/5/3 but was present in Round 8/6/4/2, then RUCLAS21 was set to RUCLAS42. If the person was not present in either Rounds 8/6/4/2 or 9/7/5/3 but was present in Round 7/5/3/1, then RUCLAS21 was set to RUCLAS31. If the person was not linked to a responding RU during any round, then RUCLAS21 was set to -15.

Geographic Variables

The round-specific variables REGION31, REGION42, REGION53, and the end-of-year status variable REGION21 indicate the Census region for the RU. REGION21 indicates the region for the 2021 portion of Round 9/7/5/3. For most analyses, REGION21 should be used.

The values and states for each region include the following:

Value	Label	States
1	Northeast	Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont

Value	Label	States
2	Midwest	Indiana, Illinois, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin
3	South	Alabama, Arkansas, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and West Virginia
4	West	Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming

Reference Period Dates

The reference period is the period of time for which data were collected in each round for each person. The reference period dates were determined during the interview for each person by the CAPI program. The round-specific beginning reference period dates are included for each person. These variables include BEGRFM31, BEGRFY31, BEGRFM42, BEGRFY42, BEGRFM53, and BEGRFY53. The reference period for Round 1 for most persons began on January 1, 2021 and ended on the date of the Round 1 interview. For RU members who joined later in Round 1, the beginning Round 1 reference date was the date the person entered the RU. For all subsequent rounds, the reference period for most persons began on the date of the previous round's interview and ended on the date of the current round's interview. Persons who joined after the previous round's interview had their beginning reference date for the round set to the day they joined the RU.

The round-specific ending reference period dates for Rounds 7/5/3/1, 8/6/4/2, and 9/7/5/3 as well as the end-of-year reference period end date variables are also included for each person. These variables include ENDRFM31, ENDRFY31, ENDRFM42, ENDRFY42, ENDRFM53, ENDRFY53, ENDRFM21, and ENDRFY21. For most persons in the sample, the date of the round's interview is the reference period end date. Note that the end date of the reference period for a person is prior to the date of the interview if the person was deceased during the round, left the RU, was institutionalized prior to that round's interview, or left the RU to join the military. For a small number of cases, the reference period dates may be recoded for confidentiality.

Reference Person Identifiers

The round-specific variables REFPRS31, REFPRS42, and REFPRS53 and the end-of-year status variable REFPRS21 identify the reference person for Rounds 7/5/3/1, 8/6/4/2, and 9/7/5/3, and as of December 31, 2021 (or the last round they were in the survey). In general, the reference person is defined as the household member 16 years of age or older who owns or rents the home. If more than one person meets this description, the household respondent identifies one from among them. If the respondent is unable to identify a person fitting this definition, the questionnaire asks for the head of household and this person is then considered the reference person for that RU. This information is collected in the Reenumeration section of the CAPI questionnaire.

Respondent Identifiers

The respondent is the person who answered the interview questions for the Reporting Unit (RU). The round-specific variables RESP31, RESP42, and RESP53 and the end-of-year status variable RESP21 identify the respondent for Rounds 7/5/3/1, 8/6/4/2, and 9/7/5/3 and as of December 31, 2021 (or the last round they were in the survey). Only one respondent is identified for each RU. In instances where the interview was completed in more than one session, only the first respondent is indicated.

There are two types of respondents. The respondent can be either an RU member or a non-RU member proxy. The round-specific variables PROXY31, PROXY42, and PROXY53 and the end-of-year status variable PROXY21 identify the type of respondent for Rounds 7/5/3/1, 8/6/4/2, 9/7/5/3 and as of December 31, 2021 (or the last round they were in the survey).

Language of Interview

The language of interview variable (INTVLANG) is a summary value of the round-specific RU-level information section question (RU30), which asks the interviewer to record the language in which the interview was completed: English, Spanish, Both English and Spanish, Other Language. Given the first round that the person was part of the study and the person's associated RU for that round, INTVLANG is assigned the interview language value reported for the person's RU for the round.

Person Status

A number of variables describe the various components reflecting each person's status for each round of data collection. These variables provide information about a person's in-scope status, Keyness status, eligibility status, and disposition status. These variables include: KEYNESS, INSCOP31, INSCOP42, INSCOP53, INSCOP21, INSC1231, INSCOPE, ELGRND31, ELGRND42, ELGRND53, ELGRND21, PSTATS31, PSTATS42, and PSTATS53. These variables are set based on sampling information and responses provided in the Reenumeration section of the CAPI questionnaire.

Through the Reenumeration section of the CAPI questionnaire, each member of an RU was classified as "Key" or "Non-Key", "in-scope" or "out-of-scope", and "eligible" or "ineligible" for MEPS data collection. To be included in the set of persons used in the derivation of MEPS person-level estimates, a person had to be a member of the civilian noninstitutionalized population for at least one day during 2021. Because a person's eligibility for the survey might have changed since the NHIS interview, a sampling reenumeration of household membership was conducted at the start of each round's interview. Only persons who were "inscope" sometime during the year, were "Key", and responded for the full period in which they were inscope were assigned positive person-level weights and thus are to be used in the derivation of person-level national estimates from the MEPS.

Note: If analysts want to subset to infants born during 2021, then newborns should be identified using $AGE21X = 0$ rather than $PSTATSxy = 51$.

Inscope

The round-specific variables INSCOP31, INSCOP42, and INSCOP53 indicate a person's in-scope status for Rounds 7/5/3/1, 8/6/4/2, and 9/7/5/3. INSCOP21, INSC1231, and INSCOPE indicate a person's in-scope status for the portion of Round 9/7/5/3 that covers 2021, the person's in-scope status as of 12/31/21, and whether a person was ever in-scope during the calendar year 2021. A person was considered as in-scope during a round or a referenced time period if they were a member of the U.S. civilian, noninstitutionalized population at some time during that round or that time period. The values of these variables taken in conjunction allow one to determine in-scope status over time (for example, becoming inscope in the middle of a round, as would be the case for newborns). These variables may contain the following values and corresponding labels:

Value	Definition
0	Incorrectly listed, or on NHIS roster but out-of-scope prior to January 1, 2021
1	Person is inscope for the whole reference period
2	Person is inscope at the start of the RU reference period, but not at the end of the RU reference period
3	Person is not inscope at the start of RU reference period, but is inscope at the end of the RU reference period. (For example, the person is inscope from the date the person joined the RU or the person was in the military in the previous round, but is no longer in the military in the current round)
4	Person is inscope during the reference period, but neither at the reference start date nor on the reference end date. (For example, person leaves an institution, goes into community, and then dies)
5	Person is out-of-scope for all of the reference period during which they are an RU member. (For example, the person is in the military)
6	Person is out-of-scope for the entire reference period and is not a member of the RU during this time period and was inscope and an RU member in an earlier round
7	Person is not in an RU, joined in a later round (or joined the RU after December 31, 2021 for INSCOP21)
8	RU Non-response and Key persons who left an RU with no tracing info and so a new RU was not formed
9	Person is not a member of an RU during this time period, and was an RU member in an earlier round

Keyness

The term "Keyness" is related to an individual's chance of being included in MEPS. A person is Key if that person is linked for sampling purposes to the set of NHIS sampled households

designated for inclusion in MEPS. Specifically, a Key person was either a member of a responding NHIS household at the time of interview, or joined a family associated with such a household after being out-of-scope at the time of the NHIS (examples of the latter situation include newborns and those returning from military service, an institution, or residence in a foreign country).

A non-Key person is one whose chance of selection for the NHIS (and MEPS) was associated with a household eligible but not sampled for the NHIS and who later became a member of a MEPS Reporting Unit. MEPS data (e.g., utilization and expenditures) were collected for the period of time a non-Key person was part of the sampled unit to provide information for family-level analyses. However, non-Key persons who leave a sample household unaccompanied by a Key, in-scope member were not followed for subsequent interviews. Non-Key individuals do not receive sample person-level weights and thus do not contribute to person-level national estimates.

The variable KEYNESS indicates a person's Keynes status. This variable is not round-specific. Instead, it is set at the time the person enters MEPS, and the person's Keynes status never changes. Once a person is determined to be Key, that person will always be Key.

It should be pointed out that a person might be Key even though not part of the civilian, noninstitutionalized portion of the U.S. population. For example, a person in the military may have been living with their civilian spouse and children in a household sampled for NHIS. The person in the military would be considered a Key person for MEPS; however, such a person would not be eligible to receive a person-level sample weight if they were never inscope during 2021.

Eligibility

The eligibility of a person for MEPS pertains to whether or not data were to be collected for that person. All of the Key in-scope persons of a sampled RU were eligible for data collection. The only non-Key persons eligible for data collection were those who happened to be living in an RU with at least one Key, in-scope person. Their eligibility continued only for the time that they were living with at least one such person. The only out-of-scope persons eligible for data collection were those who were living with Key in-scope persons, again only for the time they were living with such a person. Only military persons can meet this description (for example, a person on full-time active duty military, living with a spouse who is Key).

A person may be classified as eligible for an entire round or for some part of a round. For persons who are eligible for only part of a round (for example, persons may have been institutionalized during a round), data were collected for the period of time for which that person was classified as eligible. The round-specific variables ELGRND31, ELGRND42, and ELGRND53 and the end-of-year status variable ELGRND21 indicate a person's eligibility status for Rounds 7/5/3/1, 8/6/4/2, and 9/7/5/3 and as of December 31, 2021.

Person Disposition Status

The round-specific variables PSTATS31, PSTATS42, and PSTATS53 indicate a person’s response and eligibility status for each round of interviewing. The PSTATSxy variables indicate the reasons for either continuing or terminating data collection for each person in the MEPS. Using this variable, one could identify persons who moved during the reference period, died, were born, institutionalized, or who were in the military. Analysts should note that PSTATS53 provides a summary for all of Round 9/7/5/3, including transitions that occurred after 2021. Note that some categories may be collapsed for confidentiality purposes.

The following codes specify the value labels for the PSTATSxy variables.

Value	Definition
-1	The person was not fielded during the round or the RU was non-response
0	Incorrectly listed in RU at NHIS - applies to MEPS Round 1 only
11	Person in original RU, not full-time active military duty
12	Person in original RU, full-time active military duty, out-of-scope for whole reference period
13	Full-time student living away from home, but associated with sampled RU
14	The person is full-time active military duty during round, is inscope for part of the reference period and is in the RU at the end of the reference period
21	The person remains in a health care institution for the whole round - Rounds 8/6/4/2 and 9/7/5/3 only
22	The person leaves an institution (health care or non-health care) and rejoins the community - Rounds 8/6/4/2 and 9/7/5/3 only
23	The person leaves an institution and dies - Rounds 8/6/4/2 and 9/7/5/3 only
24	The person dies in a health care institution during the round (former RU member) - Rounds 8/6/4/2 and 9/7/5/3 only
31	Person from original RU, dies during reference period
32	Went to health care institution during reference period
33	Went to non-healthcare institution during reference period
34	Moved from original RU, outside U.S. (not as student)
35	Moved from original RU, to a military facility while on full-time active military duty
36	Went to institution (type unknown) during reference period
41	Moved from the original RU, to new RU within U.S. (new RUs include RUs originally classified as “Student RU” but which converted to “New RU”)
42	The person joins RU and is not full-time military during round

Value	Definition
43	The person's disposition as to why the person is not in the RU is unknown or the person moves and it is unknown whether the person moved inside or outside the U.S.
44	The person leaves an RU and joins an existing RU and is not both in the military and coded as inscope during the round
51	Newborn in reference period
61	Died prior to reference period (not eligible)-Round 7/5/3/1 only
62	Institutionalized prior to reference period (not eligible)-Round 7/5/3/1 only
63	Moved outside U.S., prior to reference period (not eligible)-Round 7/5/3/1 only
64	Full-time military, living on a military facility, moved prior to reference period (not eligible)-Round 7/5/3/1 only
71	Student under 24 living away at school in grades 1-12 (Non-Key)
72	Person is dropped from the RU roster as ineligible: the person is a non-Key student living away or the person is not related to reference person or the RU is the person's residence only during the school year
73	Not Key and not full-time military, moved without someone Key and in-scope (not eligible)
74	Moved as full-time military but not to a military facility and without someone Key and in-scope (not eligible this round)
81	Person moved from original RU, full-time student living away from home, did not respond

2.5.2 Navigating the MEPS Data with Information on Person Disposition Status

Since the variables PSTATS31, PSTATS42, and PSTATS53 indicate the reasons for either continuing or terminating data collection for each person in MEPS, these variables can be used to explain the beginning and ending dates for each individual's reference period of data collection, as well as which sections in the instrument each individual did not receive. By using the information included in the following table, analysts will be able to determine for each individual which sections of the MEPS questionnaire collected data elements for that person.

Some individuals have a reference period that spans an entire round, while other individuals may have data collected only for a portion of the round. When an individual's reference period does not coincide with the RU reference period, the individual's start date may be a later date, or the end date may be an earlier date, or both. In addition, some individuals have reference period information coded as "Inapplicable" (e.g., for individuals who were not actually in the household). The information in this table indicates the beginning and ending dates of reference periods for persons with various values of PSTATS31, PSTATS42, and PSTATS53. The actual dates for each individual can be found in the following variables included on this file:

BEGRFM31, BEGRFM42, BEGRFM53, BEGRFY31, BEGRFY42, BEGRFY53, ENDRFM31, ENDRFM42, ENDRFM53, ENDRFY31, ENDRFY42, ENDRFY53, ENDRFM21, and ENDRFY21.

The table below also describes the section or sections of the questionnaire that were **NOT** asked for each value of PSTATS31, PSTATS42, and PSTATS53. For example, the Priority Condition Enumeration (PE) section has questions that are not asked for deceased persons. The Closing (CL) section also contains some questions or question rosters that exclude certain persons depending on whether the person died, became institutionalized, or otherwise left the RU; however, no one is considered to have skipped the entire section. Some questions or sections (e.g., Health Status (HE), Employment (RJ, EM, EW)) are skipped if individuals are not within a certain age range. Since the PSTATS variables do not address skip patterns based on age, analysts will need to use the appropriate age variables.

The paper-and-pencil Self-Administered Questionnaire (SAQ) was designed to collect information during Panel 26 Round 2, Panel 25 Round 4, Panel 24 Round 6, and Panel 23 Round 8. A person was considered eligible to receive an SAQ if that person did not have a status of deceased or institutionalized, did not move out of the U.S. or to a military facility, was not a non-response at the time of the Round 2, Round 4, Round 6, or Round 8 interview date, and was 18 years of age or older. No RU members added in Round 3, Round 5, or Round 7 were asked to complete an SAQ questionnaire. Because PSTATS variables do not address skip patterns based on age, this questionnaire was not included in the table below. Once again, analysts will need to use the appropriate age variable, which in this case would be AGE42X. The documentation for this questionnaire appears in the SAQ section of this document under “Health Status Variables.”

Please note that the end reference date shown below for PSTATS53 reflects the Round 9/7/5/3 reference period rather than the portion of Round 9/7/5/3 that occurred during 2021.

PSTATS Value	PSTATS Description	Sections in the instrument which persons with this PSTATS value do NOT receive	Begin Reference Date	End Reference Date
-1	The person was not fielded during the round or the RU was non-response	ALL sections	Inapplicable	Inapplicable
0	Incorrectly listed in RU at NHIS - Round 1 only	ALL sections after RE	Inapplicable	Inapplicable
11	Person in original household, not FT active military duty (Person is in the same RU as the previous round)	--	PSTATS31: January 1, 2021 PSTATS42 and PSTATS53: Prior round interview date	Interview date

PSTATS Value	PSTATS Description	Sections in the instrument which persons with this PSTATS value do NOT receive	Begin Reference Date	End Reference Date
12	Person in original household, FT active military duty, out-of-scope for whole reference period.	--	PSTATS31: January 1, 2021 PSTATS42 and PSTATS53: Prior round interview date	Interview date
13	FT student living away from home, but associated with sampled household	--	PSTATS31: January 1, 2021 PSTATS42 and PSTATS53: Prior round interview date	Interview date
14	The person is FT active military duty during round and is in-scope for part of the reference period and is in the RU at the end of the reference period	--	PSTATS31: January 1, 2021 PSTATS42 and PSTATS53: Prior round interview date	PSTATS31: Interview date PSTATS42 and PSTATS53: If the person is living w/ someone Key and in-scope, then the interview date. If not living w/ someone who is Key and inscope, then the date the person joined the military
21	The person remains in a health care institution for the whole round - Rounds 8/6/4/2 and 9/7/5/3 only	All sections after RE	Inapplicable	Inapplicable
22	The person leaves a health care institution and rejoins the community - Rounds 8/6/4/2 and 9/7/5/3 only	--	Date rejoined the community	Interview date
23	The person leaves a health care institution, goes into community and then dies - Rounds 8/6/4/2 and 9/7/5/3 only	PE - Priority Conditions Enumeration HE - Health Status AC - Access to Care	Date rejoined the community	Date of Death
24	The person dies in a health care institution during the round (former household member) - Rounds 8/6/4/2 and 9/7/5/3 only	All sections after RE	Inapplicable	Inapplicable

PSTATS Value	PSTATS Description	Sections in the instrument which persons with this PSTATS value do NOT receive	Begin Reference Date	End Reference Date
31	Person from original household, dies during reference period	PE - Priority Conditions Enumeration HE - Health Status AC - Access to Care	PSTATS31: January 1, 2021 PSTATS42 and PSTATS53: Prior round interview date	Date of Death
32	Went to healthcare institution during reference period	Access to Care (AC)	PSTATS31: January 1, 2021 PSTATS42 and PSTATS53: Prior round interview date	Date institutionalized
33	Went to non-healthcare institution during reference period	Access to Care (AC)	PSTATS31: January 1, 2021 PSTATS42 and PSTATS53: Prior round interview date	Date institutionalized
34	Moved from original household, outside US	--	PSTATS31: January 1, 2021 PSTATS42 and PSTATS53: Prior round interview date	Date left the RU
35	Moved from original household, to a military facility while on FT active military duty	--	PSTATS31: January 1, 2021 PSTATS42 and PSTATS53: Prior round interview date	Date left the RU
36	Went to institution (type unknown) during reference period	Access to Care (AC)	PSTATS31: January 1, 2021 PSTATS42 and PSTATS53: Prior round interview date	Date institutionalized
41	Moved from the original household, to new household within US (new households include RUs originally classified as a student RU but which converted to a new RU; these are individuals in an RU that has split from an RU since the previous round)	--	PSTATS31: January 1, 2021 PSTATS42 and PSTATS53: Prior round interview date	Interview date

PSTATS Value	PSTATS Description	Sections in the instrument which persons with this PSTATS value do NOT receive	Begin Reference Date	End Reference Date
42	The person joins household and is not full-time military during round	--	The later date of January 1, 2021 and the date the person joined the RU	Interview date
43	The person's disposition as to why the person is not in the RU is unknown or the person moves and it is unknown whether the person moved inside or outside the U.S.	All sections after RE	Inapplicable	Inapplicable
44	The person leaves an RU and joins an existing RU and is not both in the military and coded as inscope during the round	--	PSTATS31: January 1, 2021 PSTATS42 and PSTATS53: Prior round interview date of the RU the person has joined. This may not be the interview date of the RU that the person came from	Interview date
51	Newborn in reference period	Questions where age must be > 1 Health Status (HE) Employment (RJ/EM/EW)	PSTATS31: January 1, 2021 if born prior to 2021. The date of birth if born in 2021. PSTATS42 and PSTATS53: The later of the prior round interview date and date of birth	Interview date
61	Died prior to reference period (not eligible)--Round 7/5/3/1 only	All sections after RE	Inapplicable	Inapplicable
62	Institutionalized prior to reference period (not eligible)--Round 7/5/3/1 only	All sections after RE	Inapplicable	Inapplicable
63	Moved outside U.S., prior to reference period (not eligible)--Round 7/5/3/1 only	All sections after RE	Inapplicable	Inapplicable

PSTATS Value	PSTATS Description	Sections in the instrument which persons with this PSTATS value do NOT receive	Begin Reference Date	End Reference Date
64	FT military, moved prior to reference period (not eligible)--Round 7/5/3/1 only	All sections after RE	Inapplicable	Inapplicable
71	Student under 24 living away at school in grades 1 through 12 (Non-Key)	--	PSTATS31: January 1, 2021 PSTATS42 and PSTATS53: Prior round interview date	Interview date
72	Person is dropped from the RU roster as ineligible: the person is a Non-Key student living away or the person is not related to reference person or the RU is the person's residence only during the school year	All sections after RE	Inapplicable	Inapplicable
73	Not Key and not full-time military, moved w/o someone Key and inscope (not eligible)	All sections after RE	Inapplicable	Inapplicable
74	Moved as full-time military but not to a military facility and w/o someone Key and in-scope (not eligible)	All sections after RE	Inapplicable	Inapplicable
81	Person moved from original household, FT student living away from home, did not respond	No data were collected	Inapplicable	Inapplicable

2.5.3 Demographic Variables (AGE31X-DAPID53X)

General Information

Demographic variables provide information about the demographic characteristics of each person from the MEPS HC. The characteristics include age, sex, race, ethnicity, marital status, educational attainment, and military service. As noted below, some variables have edited and imputed values. Most demographic variables on this file were asked during every round of the MEPS interview. These variables describe data for Rounds 7, 8, and 9 of Panel 23 (the panel that started in 2018); Rounds 5, 6, and 7 of Panel 24 (the panel that started in 2019); Rounds 3, 4, and 5 of Panel 25 (the panel that started in 2020); Rounds 1, 2 and 3 of Panel 26 (the panel that started in 2021); and status as of December 31, 2021.

Demographic variables that have variable names that contain ‘31’, ‘42’, or ‘53’ are round-specific variables. As mentioned in Section 2.4 “Variable Naming”, fourth year panel data for Rounds 7, 8, and 9 of Panel 23 and third year panel data for Rounds 5, 6, and 7 of Panel 24 are included in the ‘31’/‘42’/‘53’ sets of variables. For example, AGE31X represents the age data relevant to Round 7 of Panel 23, Round 5 of Panel 24, Round 3 of Panel 25 or Round 1 of Panel 26. The variable PANEL indicates the panel from which the data were derived. A value of 23 indicates Panel 23 data, a value of 24 indicates Panel 24 data, a value of 25 indicates Panel 25 data, and a value of 26 indicates Panel 26 data. The remaining demographic variables on this file are not round-specific.

The variables describing demographic status of the person as of December 31, 2021 were developed in two ways. First, the age variable (AGE21X) represents the exact age, calculated from date of birth and indicates age status as of 12/31/21. For the remaining December 31st variables [i.e., related to marital status (MARRY21X, SPOUID21, SPOUIN21), student status (FTSTU21X), and the relationship to reference persons (REFRL21X)], the following algorithm was used: data were taken from the Round 9/7/5/3 counterpart if non-missing; else, if missing, data were taken from the Round 8/6/4/2 counterpart; else from the Round 7/5/3/1 counterpart. If no valid data were available during any of these rounds of data collection, the algorithm assigned the missing value (other than -1 “Inapplicable”) from the first round that the person was part of the study. When all three rounds were set to -1, a value of -15 “Cannot be Computed” was assigned.

Age

Date of birth and age for each RU member were asked or verified during each MEPS interview (DOBMM, DOBY, AGE31X, AGE42X, AGE53X). If date of birth was available, age was calculated based on the difference between date of birth and date of interview. Inconsistencies between the calculated age and the age reported during the CAPI interview were reviewed and resolved. For purposes of confidentiality, the variables AGE31X, AGE42X, AGE53X, AGE21X, and AGELAST were top-coded at 85 years.

When date of birth was not provided but age was provided (either from the MEPS interviews or the 2017-2020 NHIS data), the month and year of birth were assigned randomly from among the possible valid options. For any cases still not accounted for, age was imputed using:

1. the mean age difference between MEPS participants with certain family relationships (where available) or
2. the mean age value for MEPS participants.

For example, a mother’s age is imputed as her child’s age plus 26, where 26 is the mean age difference between MEPS mothers and their children. A wife’s age is imputed as the husband’s age minus 3, where 3 is the mean age difference between MEPS wives and husbands. Age was imputed in this way for 19 persons on this file.

AGELAST indicates a person’s age from the last time the person was eligible for data collection during a specific calendar year. The age range for this variable is between 0 and 85.

Sex

Data on the gender of each RU member (SEX) were initially determined from the 2017 NHIS for Panel 23, from the 2018 NHIS for Panel 24, from the 2019 NHIS for Panel 25, and from the 2020 NHIS for Panel 26. The SEX variable was verified and, if necessary, corrected during each MEPS interview. The data for new RU members (persons who were not members of the RU at the time of the NHIS interviews) were also obtained during each MEPS round. When gender of the RU member was not available from the NHIS interviews and was not determined during one of the subsequent MEPS interviews, it was assigned in the following way. The person’s first name was used to assign gender if obvious (no cases were resolved in this way). If the person’s first name provided no indication of gender, then family relationships were reviewed (no cases were resolved this way). If neither of these approaches made it possible to determine the individual’s gender, gender was randomly assigned (no cases were resolved this way).

Race and Ethnicity Group

The race and the ethnic background questions were asked for each RU member during the MEPS interview. If the information was not obtained in Round 1, the questions were asked in subsequent rounds. It should be noted that race/ethnicity questions in the MEPS were revised starting with data collection in 2013 for Panel 16 Round 5, Panel 17 Round 3, and Panel 18 Round 1; this affected data starting with the FY 2012 file. Previously, there were two race questions, but starting with data collection in 2013, there is only one race question. All Asian categories listed in the second question were moved to the new single question. In addition, the new race question had additional detail for the Native Hawaiian and Other Pacific Islander categories. The main change for ethnicity is that the new questions allowed respondents to report more than one Hispanic ethnicity. Race/ethnicity data from earlier years may not be directly comparable. The following table shows the variables used for FY 2002-2011 and FY 2012-2021, with these exceptions: 1) in FY 2012, RACEV1X categories 4 and 5 were not combined but are combined starting with 2013, and 2) RACEV2X and HISPNCAT were first introduced in 2013.

MEPS Race and Ethnicity Variables, by Years, 2002 to Present

FY PUF 2002-2011	FY PUF 2012-2021
RACE	RACE
	RACEVER Used only in FY12-FY13. 1 Old Race Questions 2 New Race Questions

FY PUF 2002-2011	FY PUF 2012-2021
<p>RACEX</p> <ol style="list-style-type: none"> 1 White - No other race reported 2 Black - No other race reported 3 American Indian/Alaska Native - No other race reported 4 Asian - No other race reported 5 Native Hawaiian/Pacific Islander - No other race reported 6 Multiple races reported 	<p>RACEV1X</p> <ol style="list-style-type: none"> 1 White - No other race reported 2 Black - No other race reported 3 American Indian/Alaska Native - No other race reported 4 Asian - No other race reported (Used only in FY12. Starting in 2013, category 5 collapses into category 4.) 5 Native Hawaiian/Pacific Islander - No other race reported (Used only in FY12. Starting in 2013, category 5 collapses into category 4.) 6 Multiple races reported
	<p>RACEV2X (Starting in 2013)</p> <p>Treatment of categories 10 and 12 differed between 2013-2015 and 2016-2019</p> <ol style="list-style-type: none"> 1 White - No other race reported 2 Black - No other race reported 3 American Indian/Alaska Native - No other race reported 4 Asian Indian - No other race reported 5 Chinese - No other race reported 6 Filipino - No other race reported 10 Oth Asian/Natv Hawaiian/Pacfc Isl- No Oth 12 Multiple races reported -1 Inapplicable (Used only in FY13)
<p>RACETHNX</p> <ol style="list-style-type: none"> 1 Person is Hispanic 2 Person is Black - No other race reported/Not Hispanic 3 Person is Asian - No other race reported/Not Hispanic 4 Other race/Not Hispanic 	<p>RACETHX</p> <ol style="list-style-type: none"> 1 Hispanic 2 Non-Hispanic White only 3 Non-Hispanic Black only 4 Non-Hispanic Asian only 5 Non-Hispanic Other race or multi-race
<p>RACEAX</p> <ol style="list-style-type: none"> 1 Asian - No other race reported 2 Asian - Other race(s) reported 3 All other race assignments 	<p>RACEAX</p> <ol style="list-style-type: none"> 1 Asian - No other race reported 2 Asian - Other race(s) reported 3 All other race assignments

FY PUFIS 2002-2011	FY PUFIS 2012-2021
RACEBX 1 Black - No other race reported 2 Black - Other race(s) reported 3 All other race assignments	RACEBX 1 Black - No other race reported 2 Black - Other race(s) reported 3 All other race assignments
RACEWX 1 White - No other race reported 2 White - Other race(s) reported 3 All other race assignments	RACEWX 1 White - No other race reported 2 White - Other race(s) reported 3 All other race assignments
ETHNICITY	ETHNICITY
HISPANX 1 Hispanic 2 Not Hispanic	HISPANX 1 Hispanic 2 Not Hispanic
HISPCAT 1 Puerto Rican 2 Cuban/Cuban American 3 Dominican 4 Mexican/Mexican American 5 Central or South American 6 Non-Hispanic 91 Other Latin American 92 Other Hispanic/ Latino	HISPCAT (Used only in FY12-FY13.) 1 Puerto Rican 2 Cuban/Cuban American 3 Dominican 4 Mexican/Mexican American 5 Central or South American 6 Non-Hispanic 91 Other Latin American 92 Other Hispanic/ Latino -1 Inapplicable

FY PUFs 2002-2011	FY PUFs 2012-2021
	HISPNCAT (starting in 2013) <ul style="list-style-type: none"> 1 Mexican/Mexican American/Chicano - No other Hispanic reported 2 Puerto Rican - No other Hispanic reported 3 Cuban/Cuban American - No other Hispanic reported 4 Dominican - No other Hispanic reported 5 Central or South American - No other Hispanic reported 6 Oth Lat Am/Hisp/Latino/Spnsh orgn - No other Hispanic reported 8 Multiple Hispanic groups reported 9 Non-Hispanic -1 Inapplicable (Used only in FY13)

Race and ethnicity variables and their response categories for years prior to 2002 are available in the documentation for the FY Consolidated PUF for each data year.

Values for these variables were obtained based on the following priority order. If available, data collected were used to determine race and ethnicity. If race and/or ethnicity were not reported in the interview, then data obtained from the originally collected NHIS data were used (27 cases were resolved this way for race, and 10 cases were resolved this way for ethnicity). If still not determined, the race, and/or ethnicity were assigned based on relationship to other members of the DU using a priority ordering that gave precedence to blood relatives in the immediate family (this approach was used on 12 persons to set race and 7 persons to set ethnicity).

For the FY12 and FY13 PUFs, three new race variables were constructed for both the old and the new questions: RACEVER, RACEV1X, and RACETHX. The variable RACEVER was constructed to indicate which version of the race question(s) was asked and was included in only the 2012 and 2013 FY PUFs. RACEVER has been dropped starting with the 2014 PUF. The variables RACEV1X and RACETHX replace the variables RACEX and RACETHNX from 2002-2011. A new race variable, RACEV2X, was constructed only for the new race question and was added for the first time to the 2013 files. RACEV2X was set to -1 “Inapplicable” for persons that were not asked the new race question in FY13 only. This variable includes the expanded detail Asian categories and continues to be constructed for all PUFs.

The “multiple races reported” categories for RACEV1X and RACEV2X differ in the 2013-2015 PUFs but are the same starting with the 2016 PUF. In the 2013-2015 PUFs, persons with multiple Asian races or multiple Hawaiian/Pacific Islander races were considered multiple races for RACEV2X and were not considered multiple races for RACEV1X. Starting with the 2016

PUFs, persons with multiple Asian races or multiple Hawaiian/Pacific Islander races were no longer considered multiple races in RACEV2X.

For the FY12 and FY13 PUFs, the two Hispanic ethnicity variables from previous years were included: HISPANX and HISPCAT. The HISPANX variable continues to be constructed. The HISPCAT variable was constructed for specific Hispanic categories based only on the old question in FY12 and FY13 and HISPCAT has been dropped starting with the 2014 PUF. A new ethnicity variable, HISPNCAT, based on the new question, was introduced starting with 2013. HISPNCAT includes similar categories as HISPCAT but in a different order, and contains an additional category, 8 “Multiple Hispanic Groups Reported”, to represent any multiple responses reported. HISPNCAT was set to -1 “Inapplicable” for persons that were not asked the new ethnicity question in FY13. This variable continues to be constructed for all PUFs.

Categories have been collapsed in the variables RACEV1X, RACEV2X and HISPNCAT. For RACEV1X, new with the 2012 PUF, categories 4 and 5 were collapsed in category 4 as “ASIAN/NATV HAWAIIAN/PACFC ISL-NO OTH” starting with the 2013 PUF. For RACEV2X, new with and starting with the 2013 PUF, categories 7, 8, 9, 10, and 11 were collapsed in category 10 as “OTH ASIAN/NATV HAWAIIAN/PACFC ISL-NO OTH,” and for HISPNCAT, new with and starting with the 2013 PUF, categories 6 and 7 were collapsed in category 6 as “OTH LAT AM/HISP/LATINO/SPNSH ORGN-NO OTH”.

Language Variables: OTHLGSPK, WHTLGSPK, and HWELLSPK

Language variables (OTHLGSPK, WHTLGSPK, and HWELLSPK) were collected at the person level in the round in which the person entered the MEPS survey. Beginning with Panel 23 Round 1, the household respondent was asked for each person, age 5 or older, a person-level question to determine whether that person speaks a language other than English at home (RE1170, OTHLGSPK). If the response to OTHLGSPK was ‘yes’, then two other questions were asked. WHTLGSPK (RE1170) is a person-level question that asks whether the non-English language spoken at home is Spanish or some other language, and HWELLSPK (RE1170) is a person-level question that asks how well that person can speak English. If the response to OTHLGSPK was ‘No’, then WHTLGSPK and HWELLSPK are set to ‘-1’ (Inapplicable). Family members who are deceased or institutionalized in Round 1 are coded with a value of ‘-1’ (Inapplicable). Minors under age 5 in households have all three variables coded to “5” (Under 5 years old - Inapplicable).

For users doing multiyear analyses, please carefully review prior years’ documentation to assure correct interpretation of language variables over time.

Foreign Born Status

Three questions regarding foreign-born status were asked in the Demographic section to ascertain whether a person was born in the U.S. (RE1170), what year they came to the U.S. (RE1170) if not born in the U.S., and years lived in the U.S. (RE1170) if the response to RE1170 was ‘Don’t Know’. They replaced similar questions that had been asked in the Access to Care section prior to 2013.

The three questions were only asked once for each eligible person, that is, the first round the person was included in the interview. These new questions were asked of everyone, except deceased and institutionalized persons. The data from RE1170 are reported as the constructed variable BORNUSA. The data from RE1170 (YRCAMEUS) and RE1170 (YRSINUSA) were used to calculate the number of years a person has lived in the U.S. for the constructed variable YRSINUS. Please note that YRSINUS is a discrete variable and has collapsed categories: 1 “less than 1 year”; 2 “1 yr., less than 5 years”; 3 “5 yrs., less than 10 years”; 4 “10 yrs., less than 15 years”; 5 “15 years or more”.

Marital Status and Spouse ID

Current marital status was collected and/or updated during every round of the MEPS interview. This information was obtained in RE100 and RE1170 and is reported as MARRY31X, MARRY42X, MARRY53X, and MARRY21X. Persons under the age of 16 were coded as 6 “Under 16 - Inapplicable”. If marital status of a specified round differed from that of the previous round, then the marital status of the specified round was edited to reflect a change during the round (e.g., married in round, divorced in round, separated in round, or widowed in round).

In instances where there were discrepancies between the marital statuses of two individuals within a family, other person-level variables were reviewed to determine the edited marital status for each individual. Thus, when one spouse was reported as married and the other spouse reported as widowed, the data were reviewed to determine if one partner should be coded as 8 “Widowed in Round”.

Edits were performed to ensure some consistency across rounds. First, a person could not be coded as “Never Married” after previously being coded as any other marital status (e.g., “Widowed”). Second, a person could not be coded as “Under 16 - Inapplicable” after being previously coded as any other marital status. Third, a person could not be coded as “Married in Round” after being coded as “Married” in the round immediately preceding. Fourth, a person could not be coded as an “in Round” code (e.g., “Widowed in Round”) in two subsequent rounds. Since marital status can change across rounds and it was not feasible to edit every combination of values across rounds, unlikely sequences for marital status across the round-specific variables do exist.

The person identifier for each individual’s spouse is reported in SPOUID31, SPOUID42, SPOUID53, and SPOUID21. These are the PIDs (within each family) of the person identified as the spouse during Round 7/5/3/1, Round 8/6/4/2, and Round 9/7/5/3 and as of December 31, 2021, respectively. If no spouse was identified in the household, the variable was coded as 995 “No Spouse in House”. Those with unknown marital status are coded as 996 “Marital Status Unknown”. Persons under the age of 16 are coded as 997 “Less than 16 Years Old”.

The SPOUIN31, SPOUIN42, SPOUIN53, and SPOUIN21 variables indicate whether a person’s spouse was present in the RU during Round 7/5/3/1, Round 8/6/4/2, Round 9/7/5/3 and as of December 31, 2021 respectively. If the person had no spouse in the household, the value was

coded as 2 “Not Married/No Spouse”. For persons under the age of 16 the value was coded as 3 “Under 16 - Inapplicable”.

The SPOUID and SPOUIN variables were obtained from RE900, where the respondent was asked to identify how each pair of persons in the household was related. Analysts should note that this information was collected in a set of questions separate from the questions that asked about marital status. While editing was performed to ensure that SPOUID and SPOUIN are consistent within each round, there was no consistency check between these variables and marital status in a given round. Apparent discrepancies between marital status and spouse information may be due to any of the following causes:

1. Ambiguity as to when during a round a change in marital status occurred. This is a result of relationship information being asked for all persons living in the household at any time during the round, while marital status is asked as of the interview date (e.g., If one spouse died during the reference period, the surviving spouse’s marital status would be “Widowed in Round”, but SPOUIN and SPOUID for the same round would indicate that a spouse was present).
2. Valid discrepancies in the case of persons who are married but not living with their spouse, or separating but still living together.
3. Discrepancies that cannot be explained for either of the previous reasons.

Student Status and Educational Attainment

The variables FTSTU31X, FTSTU42X, FTSTU53X, and FTSTU21X indicate whether the person was a full-time student at the interview date (or 12/31/21 for FTSTU21X). These variables have valid values for all persons between the ages of 17 - 23 inclusive. When this question was asked during Round 1 of Panel 26, it was based on age as of the 2020 NHIS interview date.

Education questions were only asked when persons first entered MEPS, typically Round 1 for most people. It should be noted that education questions were changed with data collection in 2012 and then changed back to the original questions with data collection in 2015. The variables associated with the original education questions (data collection in 2011 and prior years and 2015 and subsequent years) are EDUCYR and HIDEG. The variable associated with the interim education question (data collection in 2012-2014) is EDUYRDEG (or EDUYRDG with collapsed categories). The variable EDRECODE relates variables for the original and interim education questions. As a result, different education variables are in the 2011-2015 PUFs based on the panel and round when a person first entered MEPS. The PUF documentation for each of the 2011-2015 years contains details about which education variables are in the respective files. Starting with the 2016 PUFs, EDUCYR and HIDEG are the only education variables on the PUFs.

EDUCYR contains the number of years of education completed when entering MEPS for individuals 5 years or older. Children under the age of 5 years were coded as -1 “Inapplicable” regardless of whether they attended school. Individuals who were 5 years of age or older and had

never attended school were coded as 0. The user should note that EDUCYR is an unedited variable and minimal data cleaning was performed on this variable.

HIDEG contains information on the highest degree of education attained at the time the individual entered MEPS. Information was obtained from three questions: highest grade completed, high school diploma, and highest degree. Persons under 16 years of age when they first entered MEPS were coded as 8 “Under 16 - Inapplicable”. In cases where the response to the highest degree question was “No Degree” and the response to the highest grade question was 13 through 17, the variable HIDEG was coded as 3 “High School Diploma”. If the response to the highest grade completed was “Refused” or “Don’t Know” and the response to the highest degree question was “No Degree”, the variable HIDEG was coded as 1 “No Degree”. The user should note that HIDEG is an unedited variable and minimal data cleaning was performed on this variable.

Military Service

Information on active duty military status was collected during each round of the MEPS interview. Persons currently on full-time active duty status are identified in the variables ACTDTY31, ACTDTY42, and ACTDTY53. Those under 16 years of age were coded as 3 “Under 16 - Inapplicable”, and those over the age of 59 were coded as 4 “Over 59 - Inapplicable”.

Relationship to the Reference Person within Reporting Units

For each Reporting Unit (RU), the person who owns or rents the DU is usually defined as the reference person. For student RUs, the student is defined as the reference person. (For additional information on reference persons, see the documentation on survey administration variables.) The relationship variables indicate the relationship of each individual to the reference person of the Reporting Unit (RU) in a given round.

For confidentiality, starting in 2013, detailed relationships were combined into more general categories in the variables REFRL31X, REFRL42X, REFRL53X, and REFRL21X. These variables replaced RFREL31X, RFREL42X, RFREL53X, and RFRELyyX used before 2013. The new and old variables are defined differently, so researchers using multiple years of MEPS should refer to prior years’ documentation to assure consistency in their data.

Note that categories for Child (4), Parent (7), and Sibling (8) for REFRL31X, REFRL42X, REFRL53X, and REFRL21X changed in 2017. In 2013-2016, these categories included biological, adoptive, step relationships, as well as in-law and foster relationships. Starting in 2017, in-law relationships are included in 91 OTHER RELATED, SPECIFY. Foster children were no longer included in MEPS starting on 2017, so those relationships no longer appear in any of the categories.

Categories	2013-2016	2017-2021
4	CHILD BIOLOGICAL/ADOPT/ IN-LAW/STEP/FOSTR	CHILD BIOLOGICAL/ADOPT/STEP
7	PARENT (BIOLGCL/ADOPT/ IN-LAW/STEP/FOSTR)	PARENT BIOLOGICAL/ADOPT/STEP
8	SIBLING (BIOLGCL/ADOPT/ IN-LAW/STEP/FOSTR)	SIBLING BIOLOGICAL/ADOPT/STEP

For the reference person, these variables have the value “Household reference person”; for all other persons in the RU, relationship to the reference person is indicated by codes representing “Spouse”, “Unmarried Partner”, “Child”, etc. A code of 91, meaning “Other Related, Specify”, was used to indicate rarely observed relationship descriptions such as “Mother of Partner”, “Partner of Sister”, etc. If the relationship of an individual to the reference person was not determined during the round-specific interview, relationships between other RU members were used, where possible, to assign a relationship to the reference person. If MEPS data from calendar year 2021 were not sufficient to identify the relationship of an individual to the reference person, relationship variables from the 2020 MEPS or NHIS data were used to assign a relationship. In the event that a meaningful value could not be determined or data were missing, the relationship variable was assigned a missing value code.

If the relationship of two individuals indicated they were spouses, but both had marital status indicating they were not married, their relationship was changed to non-marital partners. In addition, the relationship variables were edited to insure that they did not change across rounds for RUs in which the reference person did not change, with the exception of relationships identified as partner or spouse relationships.

Parent Identifiers

The variables MOPID31X, MOPID42X, MOPID53X and DAPID31X, DAPID42X, DAPID53X are round-specific and are used to identify the parents (biological, adopted, or step) of the person represented on that record. MOPID##X contains the person identifier (PID) for each individual’s mother if she lived in the RU in that panel/round of the survey, or a value of -1 “Inapplicable” if she did not. Similarly, DAPID##X contains the person identifier (PID) for each individual’s father if he lived in the RU in that panel/round of the survey, or a value of -1 “Inapplicable” if he did not. MOPID##X and DAPID##X were constructed based on information collected in the relationship grid of the instrument each round at question RE900, and include biological, adopted, and stepparents. Foster parents were not included. For persons who were not present in the household during a round, MOPID##X and DAPID##X have values of -1 “Inapplicable”.

Edits were performed to ensure that MOPID##X and DAPID##X were consistent with each individual’s age, sex, and other relationships within the family. For instance, the gender of the parent must be consistent with the indicated relationship; mothers are at least 12 years older than the person and no more than 55 years older than the person; fathers are at least 12 years older than the person; each person has no more than one mother and no more than one father; and the

PID for the person's mother and father are valid PIDs for that person's RU for the 2021 Full Year File.

2.5.4 Income and Tax Filing Variables (FILEDR21-HIEUIDX)

The file provides income and tax-related variables that were constructed primarily from data collected in the Panel 23 Round 9, Panel 24 Round 7, Panel 25 Round 5, and Panel 26 Round 3 Income sections. Person-level income amounts have been edited and imputed for every record on the full-year file, with detailed imputation flags provided as a guide to the method of editing. The tax-filing variables and some program participation variables are unedited, as discussed below.

Beginning with the income data collected for Panel 21 Round 5 and Panel 22 Round 3, two skip patterns were introduced to reduce respondent burden. Prior to Panel 21 Round 5 and Panel 22 Round 3, unemployment compensation income (IN360) was asked of all respondents eligible for income collection. Beginning with Panel 21 Round 5 and Panel 22 Round 3, IN360 was not administered for respondents who reported employment at the same current main job in the current round and the previous two rounds. Similarly, prior to Panel 21 Round 5 and Panel 22 Round 3, alimony income (IN170) was collected for all respondents eligible for income collection. Beginning with Panel 21 Round 5 and Panel 22 Round 3, IN170 was not administered for respondents who were married in the current round and married in the previous two rounds. As detailed below, weighted, sequential hot-decks were used to estimate amounts for unemployment compensation (UNEMP21X) and alimony income (ALIMP21X) for these respondents. Users should note that the introduction of these skip patterns may affect comparability of these income measures across survey years.

During imputation, logical editing and weighted, sequential hot-decks were used to estimate income amounts for missing values (both for item nonresponse and for persons in the full-year file who were not in the income rounds). Reported income components were generally left unedited (with the few exceptions noted below). Thus, analysts using these data may wish to apply additional checks for outlier values that would appear to stem from misreporting.

The editing process began with wage and salary income, WAGEP21X. Complete responses were left unedited, and this group of people was assigned WAGIMP21=1, where WAGIMP21 is the imputation flag for wage and salary data. The only exception was for a small number of persons who reported zero wage and salary income despite having been employed for pay during the year according to round-level data (see below). Data on tax filing and on taxable income sources were collected using an approach that encouraged respondents to provide information from their federal tax returns. Logical edits as well as a specific question on the wage income of the respondent (IN110) were used to assign separate income amounts to married persons whose responses were based on combined income amounts on their joint tax returns.

Persons assigned WAGIMP21=2 were those providing broad income ranges (brackets) rather than giving specific dollar amounts. Weighted sequential hot-decking was used to provide these individuals with specific dollar amounts. For this imputation, donors were persons who reported specific dollar amounts within the corresponding broad income ranges. In 2021 (as in all previous years) there were a small number of cases where WAGEP21X=0 and WAGIMP21=2.

These are cases where a married couple filing jointly reported wages with a bracket, and reported that one spouse earned \$0 of that bracketed amount. All WAGEP21X hot-deck imputations used cells defined on the basis of a conventional list of person-level characteristics including age, education, employment status, race, sex, and region.

Persons assigned WAGIMP21=3 were those who did not report wage and salary income and who were assigned WAGEP21X=0 based on not having been employed during the year.

Persons assigned WAGIMP21=4 were those who did not provide valid dollar amounts or dollar ranges, but for whom we had related information from the employment sections of the survey. In most cases this information included wages, hours, and weeks worked; for some persons, only hours and weeks worked data were reported in the employment section. The available employment section data were used to construct annualized wage amounts to be used in place of missing income section annual wage and salary data. Comparisons of reported and constructed wages and salaries using persons who provided both sorts of information yielded a high degree of confidence that employment data could be reliably used to derive values to serve in place of missing wage and salary information. To implement this approach, part-year responders were assumed to be fully employed during the remainder of the year if they were employed during the period in which they provided data. An exception was made for those who either died or were institutionalized. These persons were assigned zero wages and salaries for the time they were not in MEPS.

Hot-deck imputation was used for the remaining persons with missing WAGEP21X. Donor pools included persons whose WAGEP21X amounts were edited in the steps described above. Whenever possible, the hot-deck imputations used data on whether or not the person had been employed at any point during the year (and, if available, the number of weeks worked). Imputations for persons deemed to have been employed were conditional in nature, using only donors with positive WAGEP21X amounts (WAGIMP21=5). Imputations for WAGEP21X for the remaining persons were unconditional, using both workers and non-workers as donors (WAGIMP21=6).

After editing WAGEP21X for all persons in the full-year file, the remaining income sources were edited in the following sequence: INTRP21X, BUSNP21X, DIVDP21X, ALIMP21X, SALEP21X, TRSTP21X, PENSP21X, IRASP21X, SSECP21X, UNEMP21X, WCMPP21X, VETSP21X, CASHP21X, OTHRP21X, CHLDP21X, SSIP21X, and PUBP21X. Income components were edited sequentially, in each case using information regarding income amounts that had already been edited (so as to maintain patterns of correlation across income sources whenever possible). In all cases, bracketed responses were edited first (using hot-deck imputations from donors in corresponding brackets who gave specific dollar amounts), followed by imputations for remaining missing values. The hot-deck imputations used cells defined on the basis of income amounts already edited and a conventional list of person-level characteristics such as age, education, employment status, race, sex, and region. In addition, hot-deck imputations for CHLDP21X used family-level information concerning marital status and the number of children. Hot-deck imputations for SSIP21X and PUBP21X were also assigned using, in part, simulated program eligibility indicators that integrated state-level program eligibility criteria with data on family composition and income.

In the hot-decks for some income types, information from the National Health Interview Survey (NHIS) was used. The NHIS sample is the frame for the new sample selected for MEPS collection each year, with a year's time lag. Data from the 2017 NHIS correspond to MEPS Panel 23, data from the 2018 NHIS correspond to MEPS Panel 24, data from the 2019 NHIS correspond to MEPS Panel 25, and data from the 2020 NHIS correspond to Panel 26. Because MEPS units come from the NHIS, it is possible to match individual MEPS responding units to an NHIS unit.

Taking advantage of this matching ability, income reciprocity indicators collected by NHIS were used in imputing for missing data in certain MEPS income components - interest, dividends, business income, pensions, and Social Security. (Not all MEPS income categories have an equivalent in NHIS. Also, wage data were available from NHIS, but were not used in the MEPS imputation process.)

In cases where data on a particular income category were missing for a person in MEPS, the indicator in that income category on the NHIS file was employed, if a valid response was supplied. Indicators were examined for the entire tax-filing unit (two people in the case of married couples filing jointly; one person in all other cases).

Reported income amounts of less than one dollar were treated as missing amounts (to be hot-decked from donors with positive amounts of the corresponding income source). Also, very few cases of outlier responses were edited (primarily public sources of income that exceeded possible amounts). Otherwise, reported amounts were left unchanged.

For each income component, the corresponding xxxIMP21 variable contains an indicator concerning the method for editing/imputation. All the flag variables have the following formatted values:

- 1 = Original response used;
- 2 = Bracket converted;
- 3 = Missing value set to 0;
- 4 = Weeks worked/earnings used (WAGIMP21 only);
- 5 = Conditional hot-deck;
- 6 = Unconditional hot-deck;

Missing values were set to zero when there were too few recipients to warrant hot-deck imputations of positive values (as in the case of ALIMP21X received by males). "Conditional hot-decks" indicate instances where the respondent indicated receipt but not a specific dollar amount. In these cases, the donor pool was restricted to persons with nonzero amounts of the income source in question. "Unconditional hot-decks" indicate instances where the donor pool included persons receiving both zero and nonzero amounts (implemented in cases where there was little or no information about a person's income source).

Total person-level income (TTLP21X) is the sum of all income components with the exception of SALEP21X (to match as closely as possible the CPS definition of income; see Section 2.5.4). Some researchers may wish to define their own income measure by adding in one or both of these excluded components.

The tax variables and food stamp variables are all completely unedited. Unedited tax variables are provided to assist researchers building tax simulation programs. No efforts have been made to eliminate inconsistencies among these program participation and tax variables and other MEPS data. All of these unedited variables should be used with great care.

Income Top-Coding

All person-level income amounts on the file, including both total income and the separate sources of income, were top-coded to preserve confidentiality. For each income source, top codes were applied to the top percentile of all cases (including negative amounts that exceeded income thresholds in absolute value). In cases where less than one percent of all persons received a particular income source, all recipients were top-coded.

Top-coded income amounts were masked using a regression-based approach. The regressions relied on many of the same variables used in the hot-deck imputations, with the dependent variable in each case being the natural logarithm of the amount that the income component was in excess of its top-code threshold. Predicted values from this regression were reconverted from logarithms to levels using a smearing correction, and these predicted amounts were then added back to the top-code thresholds. This approach preserves the component-by-component weighted means (both overall and among top-coded cases), while also preserving much of the income distribution conditional on the variables contained in the regressions. At the same time, this approach ensures that every reported amount in excess of its respective threshold is altered on the public use file. The process of top-coding income amounts in this way inevitably introduces measurement error in cases where income amounts were reported correctly by respondents. Note, however, that top-coding can also help to reduce the impact of outliers that occur due to reporting errors.

Total person-level income is constructed as the sum of the adjusted person-level income components. Having constructed total income in this manner, this total was then top-coded using the same regression-based procedure described above (again masking the top percentile of cases). Finally, the components of income were scaled up or down in order to make the sources of income consistent with the newly-adjusted totals.

Poverty Status

The definitions of income, family, and poverty categories used to construct the related variables in this file were taken from the 2021 poverty statistics developed by the Current Population Survey (CPS). The categorical variable for 2021 family income as a percentage of poverty (POVCAT21) was constructed using the same method as in earlier years' files.

FAMINC21 contains total family income for each person's CPS family. Family income was derived by constructing person-level total income comprising annual earnings from wages,

salaries, bonuses, tips, commissions; business and farm gains and losses; unemployment and workers' compensation; interest and dividends; alimony, child support, and other private cash transfers; private pensions, IRA withdrawals, social security, and veterans payments; supplemental security income and cash welfare payments from public assistance, and related programs; gains or losses from estates, trusts, partnerships, S corporations, rent, and royalties; and a small amount of "other" income. Person-level income excluded tax refunds and capital gains. Person-level income totals were then summed over family members, as defined by CPSFAMID, to yield CPS family-level total income (FAMINC21).

POVLEV21 is the continuous version of the POVCAT21 variable. The POVLEV21 percentage was computed by dividing CPS family income by the applicable poverty line (based on family size and composition). POVCAT21 takes the POVLEV21 percentage for each person and classifies it into one of five poverty categories: negative or poor (less than 100%), near poor (100% to less than 125%), low income (125% to less than 200%), middle income (200% to less than 400%), and high income (greater than or equal to 400%). Persons missing CPSFAMID were treated as one-person families in constructing their poverty percentage and category.

Family income, as well as the components of person-level income, has been subjected to internal editing patterns and derivation methods that are in accordance to specific definitions, and are not being released at this time. Researchers working with a family definition other than CPSFAMID may wish to create their own versions of total family income.

Health Insurance Eligibility Units (HIEUs) are sub-family relationship units constructed to include adults plus those family members who would typically be eligible for coverage under the adults' private health insurance family plans. To construct the HIEUIDX variable, which links persons into a common HIEU, we begin with the family identification variable CPSFAMID. Working with this family ID, we define HIEUIDX using family relationships as of the end of 2021. Persons missing end-of-year relationship information are assigned to an HIEUIDX using relationship information from the last round in which they provided such information. HIEUs comprise adults, their spouses, and their unmarried natural/adoptive children age 18 and under. Prior to the 2018 data year, only opposite-sex spouses were eligible to be included in the same HIEU. Beginning with the 2018 data year, both same-sex and opposite-sex spouses are included in the same HIEU. We also include children under age 24 who are full-time students (living at home or away from home). Other children who do not live with their natural/adoptive adult parents are placed in an HIEUIDX as follows:

- Other unmarried children are placed in stepparent HIEUIDX, grandparent HIEUIDX, great-grandparent HIEUIDX, or aunt/uncle HIEUIDX.
- Children of unmarried minors are placed (along with their minor parents) in the HIEUIDX of their adult grandparents (if possible). Married minors are placed into separate HIEUs along with any spouses and children they might have.
- Some HIEUs are headed by unmarried minors, when there is no adult family member present in the CPSFAMID.

HIEUs do not, in general, comprise adult (nonmarital) partnerships, because unmarried adult partners are rarely eligible for dependent coverage under each other's insurance. The exception to this rule is that we include adult partners in the same HIEU if there is at least one (out-of-wedlock) child in the family that links to both adult partners. In cases of missing or contradictory relationship codes, HIEUs are edited by hand, with the presumption being that the adults and children form a nuclear family.

2.5.5 Person-Level Condition Variables (RTHLTH31-ADHDAGED)

Perceived Health Status

Perceived health status (RTHLTH31, RTHLTH42, and RTHLTH53) and perceived mental health status (MNHLTH31, MNHLTH42, and MNHLTH53) were collected in the Priority Conditions Enumeration (PE) section. The target persons of the questions are all current or institutionalized persons regardless of age. These questions (PE10 and PE20) asked the respondent to rate each person in the family according to the following categories: excellent, very good, good, fair, and poor.

Priority Condition Variables (HIBPDX-ADHDAGED)

The PE section was asked in its entirety in Round 1 for all current or institutionalized persons, and in Panel 26 Round 2, Panel 25 Round 4, Panel 24 Round 6, and Panel 23 Round 8 for only new RU members. In Panel 26 Round 3 and Panel 24 Round 7, the specific condition questions (except joint pain and chronic bronchitis) were asked only if the person had not reported the condition in a previous round. Also, PE questions were asked in Panel 24 Round 5 to collect calendar year data in a manner parallel to the other panels.

Priority condition variables whose names end in “DX” indicate whether the person was ever diagnosed with the condition. Chronic bronchitis, joint pain, and asthma follow-up questions (ASSTIL31/53, ASATAK31/53, and ASTHEP31/53 described below) reflect data obtained in Round 7 of Panel 23, Rounds 5 and 7 of Panel 24, Round 3 of Panel 25, and Rounds 1 and 3 of Panel 26. Diagnoses data (except attention deficit hyperactivity disorder/attention deficit disorder, diabetes, and asthma) were collected for persons over 17 years of age. If edited age is within range for the variable to be set, but the source data are missing because person’s age in CAPI is not within range, the constructed variable is set to “Cannot be Computed” (-15). Following the same pattern, attention deficit hyperactivity disorder/attention deficit disorder is asked of persons age 5 to 17, and diabetes and asthma are asked of persons of all ages. Exceptions to this pattern are the variables JTPAIN31_M18 and CHBRON31, which are described in greater detail below.

Questions were asked regarding the following conditions:

- High blood pressure, including multiple diagnoses
- Heart disease (including coronary heart disease, angina, myocardial infarction, and other unspecified heart disease)

- Stroke
- Emphysema
- Chronic bronchitis
- High cholesterol
- Cancer
- Diabetes
- Joint pain
- Arthritis
- Asthma
- Attention Deficit Hyperactivity Disorder/Attention Deficit Disorder (ADHD/ADD)

These conditions were selected because of their relatively high prevalence, and because generally accepted standards for appropriate clinical care have been developed. This information thus supplements other information on medical conditions that is gathered in other parts of the interview.

Condition data were collected at the person-by-round level (indicating if the person was ever diagnosed with the condition) and at the condition level. If the person reported having been diagnosed with a condition, the person-by-round variable was set to ‘1’ (Yes) and a condition record for that medical condition was created.

Editing of these variables focused on checking that skip patterns were consistent.

High Blood Pressure

Questions about high blood pressure (hypertension) were asked only of persons aged 18 or older. Consequently, persons aged 17 or younger were coded as “Inapplicable” (-1) on these variables. HIBPDX ascertained whether the person had ever been diagnosed as having high blood pressure (other than during pregnancy). Those who had received this diagnosis were also asked if they had been told on two or more different visits that they had high blood pressure (BPMLDX). The age of diagnosis for high blood pressure (HIBPAGED) is included in this file. This variable is top-coded to 85 years of age.

Heart Disease

Heart disease questions were asked only of persons aged 18 or older. Consequently, persons aged 17 or younger were coded as “Inapplicable” (-1) on all the variables in this set.

CHDDX - asked if the person had ever been diagnosed as having coronary heart disease

- ANGIDX - asked if the person had ever been diagnosed as having angina, or angina pectoris
- MIDX - asked if the person had ever been diagnosed as having a heart attack, or myocardial infarction
- OHRDX - asked if the person had ever been diagnosed with any other kind of heart disease or condition

The age of diagnosis for coronary heart disease (CHDAGED), angina (ANGIAGED), heart attack or myocardial infarction (MIAGED), and other kind of heart disease (OHRTAGED) are included in this file. These variables are top-coded to 85 years of age.

Respondents who answered “Yes” to a person being diagnosed with any other kind of heart disease or condition (OHRDX) were asked a follow up question (OHRRTYPE) to specify other heart diseases or conditions.

Stroke

STRKDX asked if the person (aged 18 or older) had ever been diagnosed as having had a stroke or transient ischemic attack (TIA or ministroke). Persons aged 17 or younger were coded as “Inapplicable” (-1). The age of diagnosis for stroke or TIA (STRKAGED) is included in this file. This variable is top-coded to 85 years of age.

Emphysema

EMPHDX asked if the person (aged 18 or older) had ever been diagnosed with emphysema. Persons aged 17 or younger were coded as “Inapplicable” (-1). The age of diagnosis for emphysema (EMPHAGED) is included in this file. This variable is top-coded to 85 years of age.

Chronic Bronchitis

CHBRON31 asked if the person (aged 18 or older) has had chronic bronchitis in the last 12 months. Persons aged 17 or younger were coded as “Inapplicable” (-1).

High Cholesterol

Questions about high cholesterol were asked of persons aged 18 or older. Consequently, persons aged 17 or younger were coded as “Inapplicable” (-1) on these variables. CHOLDX ascertained whether the person had ever been diagnosed as having high cholesterol. The age of diagnosis for high cholesterol (CHOLAGED) is included in this file. This variable is top-coded to 85 years of age.

Cancer

Questions about cancer were asked only of persons aged 18 or older. Consequently, persons aged 17 or younger were coded as “Inapplicable” (-1) on these variables. CANCERDX ascertained whether the person had ever been diagnosed as having cancer or a malignancy of any kind. If the respondent answered “Yes” they were asked at PE140 what type of cancer was diagnosed. CABLADDR, CABREAST, CACERVIX, CACOLON, CALUNG, CALYMPH, CAMELANO, CAMUSCLE, CAOTHER, CAPROSTA, CASKINNM, CASKINDK, and CAUTERUS indicate selection of cancer of the bladder, breast, cervix, colon, or lung; lymphoma or melanoma; cancer of the soft tissue, muscle, or fat; other type of cancer, cancer of the prostate, skin, or uterus.

Cancer of the cervix or uterus could not be reported for males, and cancer of the prostate could not be reported for females.

Recoding of Cancer Variables

Specific cancer diagnosis variables with a frequency count fewer than 20 and those considered clinically rare (i.e., appear on the National Institutes of Health’s list of rare diseases), were removed from the file for confidentiality reasons, and the corresponding variable CAOTHER, indicating diagnosis of a cancer that is not counted individually, was recoded to “Yes” (1) as necessary.

In data year 2021, the clinically rare cancers are:

- bone
- leukemia
- stomach
- brain
- liver
- testis
- esophagus
- mouth
- throat
- gallbladder
- ovary
- thyroid
- kidney
- pancreas
- larynx
- rectum

The variable CABREAST, which indicates diagnosis of breast cancer, was recoded to “Inapplicable” (-1) for males for confidentiality reasons. The corresponding value of the general cancer diagnosis variable, CANCERDX, was recoded to “Cannot be Computed” (-15), and the corresponding values of remaining specific cancer variables were recoded to “Inapplicable” (-1).

Diabetes

Prior to 2018, diabetes diagnosis was asked for each person aged 18 or older. Beginning in 2018, DIABDX_M18 replaces DIABDX where diabetes is now asked for all ages. DIABDX_M18 indicates whether each person had ever been diagnosed with diabetes (excluding gestational diabetes). The age of diagnosis for diabetes (DIABAGED) is included in this file. This variable is top-coded to 85 years of age.

Each person 18 years or older said to have received a diagnosis of diabetes was asked to complete a special self-administered questionnaire. The documentation for this questionnaire appears in the Diabetes Care Survey (DCS) section of the documentation.

Joint Pain

JTPAIN31_M18 asked if the person (aged 18 or older) had experienced pain, swelling, or stiffness around a joint in the last 12 months. This question is not intended to be used as an indicator of a diagnosis of arthritis. Persons aged 17 or younger were coded as “Inapplicable” (-1).

Joint pain questions are skipped if the person already has an arthritis condition that is specified on the conditions roster in the PE section.

Arthritis

ARTHDX asked if the person (aged 18 or older) had ever been diagnosed with arthritis. Persons aged 17 or younger were coded as “Inapplicable” (-1). Respondents who answered “Yes” were asked a follow up question to determine the type of arthritis. ARTHTYPE indicates if the diagnosis was for Rheumatoid Arthritis (1), Osteoarthritis (2), or non-specific arthritis (3). The age of diagnosis for arthritis (ARTHAGED) is included in this file and may be recoded in some cases to “Cannot be Computed” (-15) for confidentiality reasons. This variable is top-coded to 85 years of age.

Asthma

ASTHDX indicates whether a person had ever been diagnosed with asthma. The age of diagnosis for asthma (ASTHAGED) is included in this file. This variable is top-coded to 85 years of age.

Respondents who answered “Yes” to asthma diagnosis were asked additional questions. ASSTIL31 asked if the person still had asthma. ASATAK31 asked whether the person had experienced an episode of asthma or an asthma attack in the past 12 months. If the person did not experience an asthma attack in the past 12 months, a follow-up question (ASTHEP31) asked when the last asthma episode or asthma attack occurred.

Additional follow-up questions regarding asthma medication used for quick relief (ASACUT31), preventive medicine (ASPREV31), and peak flow meters (ASPKFL31) were asked. These questions were asked if the person reported having been diagnosed with asthma (ASTHDX = 1). ASACUT31 asked whether, during the last three months, the person had used the kind of prescription inhaler “that you breathe in through your mouth” that gives quick relief from asthma symptoms. ASPREV31 asked whether the person had ever taken the preventive kind of asthma medicine used every day to protect the lungs and prevent attacks, including both oral medicine and inhalers. ASPKFL31 indicates whether the person with asthma had a peak flow meter at home.

Respondents who answered “Yes” to ASACUT31 were asked whether the person had used more than three canisters of this type of inhaler in the past three months (ASMRCN31). Respondents who answered “Yes” to ASPREV31 were asked whether the person now took this kind of medication daily or almost daily (ASDALY31). Respondents who answered “Yes” to ASPKFL31 were asked if the person ever used the peak flow meter (ASEVFL31). Those respondents who answered “Yes” to ASEVFL31 were asked when the person last used the peak flow meter (ASWNFL31).

Beginning in 2018, questions regarding asthma medication used for quick relief, preventive medicine, and peak flow meters are now implemented starting with Panel 22 Round 3 and Panel 23 Round 1. With the extension of rounds and an additional panel beginning in 2020, Round 5/3 asthma variables have been added. The asthma variables included in this file are:

ASSTIL31 (Does Person Still Have Asthma - RD 3/1)

ASSTIL53 (Does Person Still Have Asthma - RD 5/3)

ASATAK31 (Asthma Attack Last 12 Mos - RD 3/1)

ASATAK53 (Asthma Attack Last 12 Mos - RD 5/3)
ASTHEP31 (When Was Last Episode of Asthma - RD 3/1)
ASTHEP53 (When Was Last Episode of Asthma - RD 5/3)
ASACUT31 (Used Acute Pres Inhaler Last 3 Mos- RD 3/1)
ASACUT53 (Used Acute Pres Inhaler Last 3 Mos- RD 5/3)
ASPREV31 (Ever Used Prev Daily Asthma Meds - RD 3/1)
ASPREV53 (Ever Used Prev Daily Asthma Meds - RD 5/3)
ASPKFL31 (Have Peak Flow Meter at Home - RD 3/1)
ASPKFL53 (Have Peak Flow Meter at Home - RD 5/3)
ASMRCN31 (Used >3 Acute Cn Pres Inh Last 3 Mos - RD 3/1)
ASMRCN53 (Used >3 Acute Cn Pres Inh Last 3 Mos - RD 5/3)
ASDALY31 (Now Take Prev Daily Asthma Meds - RD 3/1)
ASDALY53 (Now Take Prev Daily Asthma Meds - RD 5/3)
ASEVFL31 (Ever Used Peak Flow Meter - RD 3/1)
ASEVFL53 (Ever Used Peak Flow Meter - RD 5/3)
ASWNFL31 (When Last Used Peak Flow Meter - RD 3/1)
ASWNFL53 (When Last Used Peak Flow Meter - RD 5/3)

It may appear that there are discrepancies between the diagnosis variable and the follow-up variables. If a person reported asthma in the PE section in Panel 26 Round 3, ASATAK31 and ASSTIL31 will be set to “Inapplicable” (-1) as the person had not reported asthma in Round 1.

Attention Deficit Hyperactivity Disorder/Attention Deficit Disorder

ADHDADDX asked if persons aged 5 through 17 had ever been diagnosed as having Attention Deficit Hyperactivity Disorder or Attention Deficit Disorder. Persons younger than 5 or older than 17 were coded as “Inapplicable” (-1). The age of diagnosis for attention deficit hyperactivity disorder/attention deficit disorder (ADHDAGED) is included in this file.

2.5.6 Health Status Variables (IADLHP31-ADOVER42)

Due to the overlapping panel design of the MEPS (Round 7 for Panel 23, Round 5 for Panel 24, Round 3 for Panel 25 and Round 1 for Panel 26 overlapped; Round 8 for Panel 23, Round 6 for

Panel 24, Round 4 for Panel 25, and Round 2 for Panel 26 overlapped; and Round 7 for Panel 24 and Round 3 for Panel 26 overlapped), data from overlapping rounds have been combined across panels. In 2020, data collection was expanded beyond five rounds. In 2021, variables ending in “31” reflect data obtained in Round 7 of Panel 23, Round 5 of Panel 24, Round 3 of Panel 25 and Round 1 of Panel 26. Variables ending in “42” reflect data obtained in Round 8 of Panel 23, Round 6 of Panel 24, Round 4 of Panel 25 and Round 2 of Panel 26. Variables ending in “53” reflect data obtained in Round 7 of Panel 24 and Round 3 of Panel 26. Health Status variables whose names end in “21” indicate a full-year measurement. For persons in Panel 23, Round 7 extended from 2020 into 2021; for persons in Panel 24, Round 5 extended from 2020 into 2021; and for persons in Panel 25, Round 3 extended from 2020 into 2021. Therefore, for these people, some information from late 2020 is included for variables that have names ending in “31”.

Health Status variables in this data release can be classified into several conceptually distinct sets:

- IADL (Instrumental Activities of Daily Living) and ADL (Activities of Daily Living) limitations
- Functional and activity limitations
- Hearing, vision problems
- Disability status
- Any limitations
- Child health and preventive care
- Additional health
- Self-administered questionnaire
- Diabetes care survey

In general, Health Status variables involved the construction of person-level variables based on information collected in the Health Status section of the questionnaire. Many Health Status questions were initially asked at the family level to ascertain if anyone in the household had a particular problem or limitation. These were followed up with questions to determine which household member had each problem or limitation. All information ascertained at the family level has been brought to the person level for this file. Logical edits were performed in constructing the person-level variables to assure that family-level and person-level values were consistent. Particular attention was given to cases where missing values were reported at the family level to ensure that appropriate information was carried to the person level.

Inapplicable cases occurred when a question was never asked because of a skip pattern in the survey (e.g., some follow-up verification questions were not asked about individuals who were 13 years of age or older; questions pertaining to children’s health status were not asked about

individuals older than 17). Inapplicable cases are coded as -1. In addition, deceased persons were coded as “Inapplicable” (-1).

Each of the sets of variables listed above will be described in turn.

IADL and ADL Limitations

IADL Help

The Instrumental Activities of Daily Living (IADL) Help or Supervision variable IADLHP31 was constructed from a series of three questions administered in the Health Status section of the interview in Panel 23 Round 7, Panel 24 Round 5, Panel 25 Round 3 and Panel 26 Round 1. In 2021, the IADL questions were also administered in Panel 24 Round 7 and Panel 26 Round 3 and the variable IADLHP53 is included in this file.

The initial question (HE10) determined if anyone in the family received help or supervision with IADLs such as using the telephone, paying bills, taking medications, preparing light meals, doing laundry, or going shopping. If the response was “Yes”, a follow-up question (HE20) was asked to determine which household member(s) received this help or supervision. For persons under age 13, a final verification question (HE30) was asked to confirm that the IADL help or supervision was the result of an impairment or physical or mental health problem. If the response to the final verification question was “No”, IADLHP31 or IADLHP53 was coded “No” for persons under the age of 13.

If no one in the family was identified as receiving help or supervision with IADLs, all members of the family were coded as receiving no IADL help or supervision. In cases where the response to the family-level question was “Refused” (-7), “Don’t Know” (-8), or “Cannot be Computed” (-15), all persons were coded according to the family-level response. In cases where the response to the family-level question (HE10) was “Yes” but no specific individuals were identified in the follow-up question as having IADL difficulties, all persons were coded as “Don’t Know” (-8).

ADL Help

The Activities of Daily Living (ADL) Help or Supervision variable ADLHLP31 was constructed in the same manner, and for the same persons, as the IADL help variable, but using questions HE40-HE60 in Panel 23 Round 7, Panel 24 Round 5, Panel 25 Round 3 and Panel 26 Round 1. In 2021, the ADL questions were also administered in Panel 24 Round 7 and Panel 26 Round 3 and the variable ADLHLP53 is included in this file. Coding conventions for missing data were the same as for the IADL variable.

Functional and Activity Limitations

A series of health status questions was asked related to functional limitations; use of assistive technology and social/recreational limitations; work, housework, and school limitations; and cognitive limitations. The ‘31’ versions of these variables incorporate data collected in Panel 23 Round 7, Panel 24 Round 5, Panel 25 Round 3 and Panel 26 Round 1.

The ‘53’ versions of these variables incorporate data collected in Panel 24 Round 7 and Panel 26 Round 3.

Functional Limitations

A series of questions was asked that pertained to functional limitations, which are defined as difficulty in performing certain specific physical actions. WLKLIM31/53 was the filter question. These variables were derived from a question (HE90) that was asked at the family level: “Does anyone in the family have difficulties walking, climbing stairs, grasping objects, reaching overhead, lifting, bending or stooping, or standing for long periods of time?” If the answer was “No”, then all family members were coded as “No” (2) on WLKLIM31/53. If the answer was “Yes”, then the specific persons who had any of these difficulties were identified and coded as “Yes” (1), and remaining family members were coded as “No” (2). If the response to the family-level question was “Don’t Know” (-8), “Refused” (-7), “Cannot be Computed” (-15), or “Inapplicable” (-1), then the corresponding missing value code was applied to each family member’s value for WLKLIM31/53. If the answer to HE90 was “Yes” (1) but no specific individual was named as experiencing such difficulties, then each family member was assigned “Don’t Know” (-8). Deceased persons were assigned a code of “Inapplicable” (-1) for WLKLIM31/53.

If WLKLIM31/53 was coded “Yes” (1) for any family member, a subsequent series of questions was administered. The series of questions for which WLKLIM31/53 served as a filter is as follows:

- LFTDIF31/53 - difficulty lifting 10 pounds
- STPDIF31/53 - difficulty walking up 10 steps
- WLKDIF31/53 - difficulty walking 3 blocks
- MILDIF31/53 - difficulty walking a mile
- STNDIF31/53 - difficulty standing 20 minutes
- BENDIF31/53 - difficulty bending or stooping
- RCHDIF31/53 - difficulty reaching over head
- FNGRDF31/53 - difficulty using fingers to grasp

This series of questions was asked separately for each person whose response to WLKLIM31/53 was coded “Yes” (1). The series of questions was not asked for other individual family members whose response to WLKLIM31/53 was “No” (2). In addition, this series was not asked about family members who were less than 13 years of age, regardless of their status on WLKLIM31/53. These questions were not asked about deceased family members. In such cases (i.e., WLKLIM31/53 = 2, or age < 13, or PSTATS31/53 = 23, 24, or 31), each question in the series was coded as “Inapplicable” (-1). Finally, if responses to WLKLIM31/53 were “Refused” (-7), “Don’t Know” (-8), “Cannot be Computed” (-15), or otherwise “Inapplicable” (-1), then each question in this series was coded as “Inapplicable” (-1).

Analysts should note that WLKLIM31/53 was asked of all household members, regardless of age. For the subsequent series of questions, however, persons less than 13 years old were skipped

and coded as “Inapplicable” (-1). Therefore, it is possible for someone age 12 or younger to have a code of “Yes” (1) on WLKLIM31/53, and also to have codes of “Inapplicable” on the subsequent series of questions.

Use of Assistive Technology and Social/Recreational Limitations

The variables indicating use of assistive technology (AIDHLP31/53 from question HE70) and social/recreational limitations (SOCLIM31/53, from question HE230) were collected initially at the family level. If there was a “Yes” (1) response to the family-level question, a second question identified the specific individual(s) to whom the “Yes” response pertained. Each individual identified as having the difficulty was coded “Yes” (1) for the appropriate variable; all remaining family members were coded “No” (2). If the family-level response was “Refused” (-7), “Don’t Know” (-8), or “Cannot be Computed” (-15), all persons were coded with the family-level response. In cases where the family-level response was “Yes” but no specific individual was identified as having difficulty, all family members were coded as “Don’t Know” (-8).

Work, Housework, and School Limitations

The variable indicating any limitation in work, housework, or school (ACTLIM31/53) was constructed using questions HE190-HE200. Specifically, information was collected initially at the family level. If there was a “Yes” (1) response to the family-level question (HE190), a second question (HE200) identified the specific individual(s) to whom the “Yes” (1) response pertained. Each individual identified as having a limitation was coded “Yes” (1) for the appropriate variable; all remaining family members were coded “No” (2). If the family-level response was “Refused” (-7), “Don’t Know” (-8), or “Cannot be Computed” (-15), all persons were coded with the family-level response. In cases where the family-level response was “Yes” (1) but no specific individual was identified as having limitation, all family members were coded as “Don’t Know” (-8). Persons less than five years old were coded as “Inapplicable” (-1) on ACTLIM31/53.

If ACTLIM31/53 was “Yes” (1) and the person was 5 years of age or older, a follow-up question (HE210) was asked to identify the specific limitation or limitations for each person. These included working at a job (WRKLIM31/53), doing housework (HSELIM31/53), or going to school (SCHLIM31/53). Respondents could answer “Yes” (1) or “No” (2) to each activity; thus a person could report limitations in multiple activities. WRKLIM31/53, HSELIM31/53, and SCHLIM31/53 have values of “Yes” (1) or “No” (2) only if ACTLIM31/53 was “Yes” (1); each variable was coded as “Inapplicable” (-1) if ACTLIM31/53 was “No” (2). When ACTLIM31/53 was “Refused” (-7), these variables were all coded as “Refused” (-7); when ACTLIM31/53 was “Don’t Know” (-8), these variables were all coded as “Don’t Know” (-8); and when ACTLIM31/53 was “Cannot be Computed” (-15), these variables were all coded as “Cannot be Computed” (-15). If a person was under 5 years old or was deceased, WRKLIM31/53, HSELIM31/53, and SCHLIM31/53 were each coded as “Inapplicable” (-1).

An additional question (UNABLE31/53) asked if the person was completely unable to work at a job, do housework, or go to school. Those persons who were coded “No” (2), “Refused” (-7), “Don’t Know” (-8), or “Cannot be Computed” (-15) on ACTLIM31/53, were under 5 years of age, or were deceased were coded as “Inapplicable” (-1) on UNABLE31/53. UNABLE31/53 was asked once for whichever set of WRKLIM31/53, HSELIM31/53, and SCHLIM31/53 the person had limitations; if a person was limited in more than one of these three activities, UNABLE31/53

did **not** specify if the person was completely unable to perform all of them, or only some of them.

Cognitive Limitations

The variable indicating any cognitive limitation (COGLIM31/53) was collected at the family level as a three-part question (HE250A to HE250C), asking if any of the adults in the family (1) experience confusion or memory loss, (2) have problems making decisions, or (3) require supervision for their own safety. If a “Yes” response was obtained to any item, the persons affected were identified in HE260, and COGLIM31/53 was coded as “Yes” (1). Remaining family members not identified were coded as “No” (2) for COGLIM31/53.

If responses to HE250A through HE250C were all “No”, or if two of three were “No” (2) and the remaining was “Refused” (-7), “Don’t Know” (-8), or “Cannot be Computed” (-15), all family members were coded as “No” (2). If responses to the three questions were combinations of “Don’t Know” (-8), “Refused” (-7), and missing, all persons were coded as “Don’t Know” (-8). If the response to any of the three questions was “Yes” (1) but no individual was identified in HE260, all persons were coded as “Don’t Know” (-8).

COGLIM31/53 reflects whether **any** of the three component questions is “Yes” (1). Family members with one, two, or three specific cognitive limitations cannot be distinguished. In addition, because the question asked specifically about adult family members, all persons less than 18 years of age are coded as “Inapplicable” (-1) on this question.

Hearing, Vision Problems

A series of questions (HE270 to HE310), asked in Panel 23 Round 8, Panel 24 Round 6, Panel 25 Round 4, and Panel 26 Round 2, provides information on hearing and visual impairment. Household members less than one year old and deceased RU members were coded as “Inapplicable” (-1).

The hearing impairment variable, DFHEAR42, indicates whether a person has serious difficulty hearing. This variable was based on two questions, HE270 and HE280. The initial question (HE270) determined if anyone in the family had difficulty hearing. If the response was “Yes” (1), a follow-up question (HE280) was asked to determine which household member(s) had a hearing impairment. If the family-level response was “Don’t Know” (-8), “Refused” (-7), or “Cannot be Computed” (-15), all persons were coded with the family-level response. In cases where the family-level response was “Yes” (1) but no specific individual was identified as having serious difficulty hearing, all family members were coded as “Don’t Know” (-8).

The visual impairment variable, DFSEE42, indicates whether a person has serious difficulty seeing. This variable was based on two questions, HE290C and HE300. The initial question (HE290C) determined if anyone in the family had difficulty seeing. If the response was “Yes” (1), a follow-up question (HE300) was asked to determine which household member(s) had a seeing impairment. If the family-level response was “Don’t Know” (-8), “Refused” (-7), or “Cannot be Computed” (-15), all persons were coded with the family-level response. In cases where the family-level response was “Yes” (1) but no specific individual was identified as having serious difficulty seeing, all family members were coded as “Don’t Know” (-8).

Disability Status

A series of questions (HE310 to HE380) in Panel 23 Round 8, Panel 24 Round 6, Panel 25 Round 4, and Panel 26 Round 2 provides information on cognitive difficulty, difficulty walking or climbing stairs, difficulty dressing or bathing, and difficulty doing errands. Questions regarding cognitive difficulty, difficulty walking or climbing stairs, and difficulty dressing or bathing were asked of household members 5 years of age and older. The question regarding difficulty doing errands was asked of household members 15 years of age and older. Deceased RU members were coded “Inapplicable” (-1).

DFCOG42 indicates whether a person had serious cognitive difficulty. This variable was based on two questions, HE310 and HE320. The initial question (HE310) determined if anyone in the family had difficulty concentrating, remembering or making decisions. If the response was “Yes” (1), a follow-up question (HE320) was asked to determine which household member(s) had difficulty concentrating, remembering or making decisions. If the family-level response was “Don’t Know” (-8), “Refused” (-7), or “Cannot be Computed” (-15), all persons were coded with the family-level response. In cases where the family-level response was “Yes” (1) but no specific individual was identified as having serious cognitive difficulty, all family members were coded as “Don’t Know” (-8).

DFWLKC42 indicates whether a person has serious difficulty walking or climbing stairs. This variable was based on two questions, HE330 and HE340. The initial question (HE330) determined if anyone in the family had serious difficulty walking or climbing stairs. If the response was “Yes” (1), a follow-up question (HE340) was asked to determine which household member(s) had difficulty walking or climbing stairs. If the family-level response was “Don’t Know” (-8), “Refused” (-7), or “Cannot be Computed” (-15), all persons were coded with the family-level response. In cases where the family-level response was “Yes” (1) but no specific individual was identified as having serious difficulty walking or climbing stairs, all family members were coded as “Don’t Know” (-8).

DFDRSB42 indicates whether a person has difficulty dressing or bathing. This variable was based on two questions, HE350 and HE360. The initial question (HE350) determined if anyone in the family had difficulty dressing or bathing. If the response was “Yes” (1), a follow-up question (HE360) was asked to determine which household member(s) had difficulty dressing or bathing. If the family-level response was “Don’t Know” (-8), “Refused” (-7), or “Cannot be Computed” (-15), all persons were coded with the family-level response. In cases where the family-level response was “Yes” (1) but no specific individual was identified as having difficulty dressing or bathing, all family members were coded as “Don’t Know” (-8).

DFERN42 indicates whether a person has difficulty doing errands alone. This variable was based on two questions, HE370 and HE380. The initial question (HE370) determined if anyone in the family had difficulty doing errands alone. If the response was “Yes” (1), a follow-up question (HE380) was asked to determine which household member(s) had difficulty doing errands alone. If the family-level response was “Don’t Know” (-8), “Refused” (-7), or “Cannot be Computed” (-15), all persons were coded with the family-level response. In cases where the family-level response was “Yes” (1) but no specific individual was identified as having difficulty doing errands alone, all family members were coded as “Don’t Know” (-8).

Any Limitation Rounds 7 and 8 (Panel 23) / Rounds 5, 6, and 7 (Panel 24) / Rounds 3 and 4 (Panel 25) / Rounds 1, 2, and 3 (Panel 26)

ANYLMI21 summarizes whether a person had any IADL, ADL, functional, or activity limitations in any of the pertinent rounds. ANYLMI21 was built using the component variables IADLHP31/53, ADLHLP31/53, WLKLIM31/53, ACTLIM31/53, DFSEE42, and DFHEAR42. If any of these components was coded “Yes”, then ANYLMI21 was coded “Yes” (1). If all components were coded “No”, then ANYLMI21 was coded “No” (2). If all the components were “Inapplicable” (-1), then ANYLMI21 was coded as “Inapplicable” (-1). If all the components had missing value codes (i.e., -7, -8, or -1), ANYLMI21 was coded as “Cannot be Computed” (-15). If some components were “No” and others had missing value codes, ANYLMI21 was coded as “Cannot be Computed” (-15). The exception to this latter rule was for children younger than five years old, who were not asked questions that are the basis for ACTLIM31/53; for these RU members, if all other components were “No”, then ANYLMI21 was coded as “No” (2). The variable label for ANYLMI21 departs slightly from conventions. Typically, variables that end in “21” refer only to 2021. However, some of the variables used to construct ANYLMI21 were assessed in 2022, so some information from early 2022 is incorporated into this variable.

Child Health and Preventive Care

Questions were asked about each child (under the age of 18 excluding deceased children) in the applicable age subgroups to which they pertained. For the Child Supplement variables, a code of “Inapplicable” (-1) was assigned if a person was deceased, was not in the appropriate Round 2, 4, 6, or 8 or was not in the applicable age subgroup as of the interview date. This public use dataset contains variables and frequency distributions from the Child Health and Preventive Care (CS) Section associated with 7,519 children who were eligible for the CS section. Children were eligible for this section when PSTATS42 was not equal to 23, 24, 31 (Deceased) and $0 \leq \text{AGE42X} \leq 17$. Of these children, 5,379 were assigned a positive person-level weight for 2021 ($\text{PERWT21F} > 0$). Cases not eligible for the CS section should be excluded from estimates made with the CS section.

Starting in 2018, the Consumer Assessment of Healthcare Providers and Systems (CAHPS) and Columbia Impairment Scale (CIS) series of questions will be administered every other year. CAPI will administer the CAHPS and CIS series as follows:

- In Rounds 2 and 6 for panels whose Round 1 collection occurred in an odd year
- In Rounds 4 and 8 for panels whose Round 1 collection occurred in an even year

Therefore, since Panel 26 Round 1 collection started in 2021, Panel 25 Round 1 collection started in 2020, Panel 24 Round 1 collection started in 2019, and Panel 23 Round 1 started in 2018, the CAHPS and CIS questions were asked, and these variables are included in the 2021 dataset.

In addition, the child preventive care series will be administered every other year beginning in 2018. CAPI will administer the child preventive care series as follows:

- In Rounds 2 and 6 for panels whose Round 1 collection occurred in an even year
- In Rounds 4 and 8 for panels whose Round 1 collection occurred in an odd year

Therefore, the child preventive care questions were not asked in 2021 and are not included in the 2021 dataset.

The following table summarizes the data years in which the subsections of CS section are asked.

Child Health Survey Section	2017	2018	2019	2020	2021
Special Health Care Needs	X	X	X	X	X
Child Preventive Care	X	X		X	
Columbia Impairment Scale (CIS)	X		X		X
Consumer Assessment of Healthcare Providers and Systems (CAHPS)	X		X		X

Children with Special Health Care Needs Screener (ages 0 - 17)

The Children with Special Health Care Needs (CSHCN) Screener instrument was developed through a national collaborative process as part of the Child and Adolescent Health Measurement Initiative (CAHMI) coordinated by the Foundation for Accountability. A key reference for this screener instrument is Bethel et al. (2002).

These questions are asked about children ages 0-17. In general, the CSHCN screener identifies children with activity limitation or need or use of more health care or other services than is usual for most children of the same age. When a response to a gate question was set to “No” (2), “Refused” (-7), “Don’t Know” (-8), or “Cannot be Computed” (-15), follow-up variables based on the gate question were coded as “Inapplicable” (-1).

The variable CSHCN42 identifies children with special health care needs, and was created using the CSHCN screener questions according to the specifications in the reference above. The CSHCN screener questions consist of a series of question sequences about the following five health consequences: the need or use of medicines prescribed by a doctor; the need or use of more medical care, mental health, or education services than is usual for most children; being limited or prevented in doing things most children can do; the need or use of special therapy such as physical, occupational, or speech therapy; and the need or use of treatment or counseling for emotional, developmental, or behavioral problems. Parents who responded “yes” to any of the “initial” questions in the five question sequences were then asked to respond to up to two follow-up questions about whether the health consequence was attributable to a medical, behavioral, or other health condition lasting or expected to last at least 12 months. Children with positive responses to at least one of the five health consequences along with all of the follow-up questions

were identified as having a Special Health Care Need. Children with a “no” response for all five question sequences were considered NOT to have a Special Health Care Need. Those children whose “special health care need” status could not be determined (due to missing data for any of the questions) were coded as “Unknown” for CSHCN42. More [information about the CSHCN screener questions](#) can be obtained from the website for the Child and Adolescent Health Measurement Initiative.

The CSHCN screener questions were:

- CHPMED42 - child needs or uses prescribed medicines
- CHPMHB42 - prescribed medicines were because of a medical, behavioral, or other health condition
- CHPMC42 - health condition that causes a person to need prescribed medicines has lasted or is expected to last for at least 12 months
- CHSERV42 - child needs or uses more medical care, mental health, or education services than is usual for most children of the same age
- CHSRHB42 - child needs or uses more medical and other service because of a medical, behavioral, or other health condition
- CHSRC42 - health condition that causes a person to need or use more medical and other services has lasted or is expected to last for at least 12 months
- CHLIMI42 - child is limited or prevented in any way in ability to do the things most children of the same age can do
- CHLIHB42 - child is limited in the ability to do the things most children can do because of a medical, behavioral, or other health condition
- CHLICO42 - health condition that causes a person to be limited in the ability to do the things most children can do has lasted or is expected to last for at least 12 months
- CHTHER42 - child needs or gets special therapy such as physical, occupational, or speech therapy
- CHTHHB42 - child needs or gets special therapy because of a medical, behavioral, or other health condition
- CHTHCO42 - health condition that causes a person to need or get special therapy has lasted or is expected to last for at least 12 months
- CHCOUN42 - child has an emotional, developmental, or behavioral problem for which they need or get treatment or counseling

CHEMPB42 - problem for which a person needs or gets treatment or counseling is a condition that has lasted or is expected to last for at least 12 months

CSHCN42 - identifies children with special health care needs

Columbia Impairment Scale (ages 5 - 17) (included in alternating years only)

These questions inquired about possible child behavioral problems and were asked in previous years. Respondents were asked to rate on a scale from 0 to 4, where “0” indicates “No Problem” and “4” indicates “A Very Big Problem”, how much of a problem the child has with thirteen specified activities. A key reference for the Columbia Impairment Scale is Bird et al. (1996).

Certain questions in this series were coded to “Asked, but Inapplicable” (99) when the question was not applicable for a specific child. For example, if a child’s mother was deceased, a question about how much of a problem a child has getting along with his/her mother would be set to “Asked, but Inapplicable” (99). Similarly, the question about problems getting along with siblings would be set to “Asked, but Inapplicable” (99) for children with no siblings. Variables in this set include:

GETTRB42 - problem with getting into trouble

MOMPRO42 - getting along with mother

DADPRO42 - getting along with father

UNHAP42 - feeling unhappy or sad

SCHLBH42 - (his/her) behavior at school

HAVFUN42 - having fun

ADUPRO42 - getting along with adults

NERVAF42 - feeling nervous or afraid

SIBPRO42 - along with brothers and sisters

KIDPRO42 - getting along with other kids

SPRPRO42 - getting involved in activities like sports or hobbies

SCHPRO42 - (his/her) schoolwork

HOMEBH42 - (his/her) behavior at home

CAHPS® (Consumer Assessment of Healthcare Providers and Systems) ages 0 - 17 (included in alternating years only)

The health care quality measures were taken from the health plan version of CAHPS®, an AHRQ-sponsored family of survey instruments designed to measure quality of care from the

consumer's perspective. All of the CAHPS® variables refer to events experienced in the last 12 months. The variables included from the CAHPS® are:

- CHILCR42 - whether a person had an illness, injury, or condition that needed care right away from a clinic, emergency room, or doctor's office
- CHILWW42 - how often a person got care as soon as was needed (coded as "-1 Inapplicable" when CHILCR42 = 2, -7, -8, or -15)
- CHRTCR42 - whether any appointments were made
- CHRTWW42 - how often a person got an appointment for health care as soon as was needed (coded as "-1 Inapplicable" when CHRTCR42 = 2, -7, -8, or -15)
- CHAPPT42 - how many times a person went to a doctor's office or clinic for health care
- CHLIST42 - how often a person's doctors or other health providers listened carefully to the parent (coded as "-1 Inapplicable" when CHAPPT42 = 0, -7, -8, or -15)
- CHEXPL42 - how often a person's doctors or other health providers explained things in a way the parent could understand (coded as "-1 Inapplicable" when CHAPPT42 = 0, -7, -8, or -15)
- CHRESP42 - how often a person's doctors or other health providers showed respect for what the parent had to say (coded as "-1 Inapplicable" when CHAPPT42 = 0, -7, -8, or -15)
- CHPRTM42 - how often doctors or other health providers spent enough time with a person (coded as "-1 Inapplicable" when CHAPPT42 = 0, -7, -8, or -15)
- CHHECR42 - rating of health care from 0 to 10 where 0 = Worst health care possible and 10 = Best health care possible (coded as "-1 Inapplicable" when CHAPPT42 = 0, -7, -8, or -15)
- CHSPEC42_M18 - whether a person made an appointment to see a specialist
- CHEYRE42_M18 - how often did a person get appointments to see a specialist (coded as "-1 Inapplicable" when CHSPEC42 = 2, -7, -8, or -15)

Additional Health Variables

LSTETH53 (has person lost all natural (permanent) teeth), PHYEXE53 (currently spends half hour or more in moderate to vigorous physical activity at least five times a week), and OFTSMK53 (how often smoke cigarettes) are asked in the Additional Healthcare Questions (AH) section. These questions are asked every year of each person 18 years or older. A code of "Inapplicable" (-1) was assigned if the person was deceased or less than 18 years old. In 2021,

these variables include data collected in Panel 23 Round 9, Panel 24 Round 7, Panel 25 Round 5, and Panel 26 Round 3.

2021 Self-Administered Questionnaires

The MEPS distributes several hard-copy, self-administered questionnaires (SAQs) to collect health-related information from different subpopulations of MEPS participants. The Diabetes Care Survey is distributed every year, while other SAQs are distributed only in select years. The table below lists the SAQs distributed in select years and the years they are collected, while the remainder of this section describes in detail the SAQs collected in the current data year.

Self-Administered Questionnaire	2018	2019	2020	2021
SAQ		X		X
Preventive SAQ (PSAQ)	X		X	
Mental Health SAQ (MSAQ), embedded in PSAQ			X	
Veteran’s SAQ (VSAQ)	X	X		
Social Determinants of Health (SDOH)				X

Self-Administered Questionnaire (SAQ)

The Self-Administered Questionnaire (SAQ) is a paper-and-pencil questionnaire that includes core questions about health status, health care quality, and preventive health care measures of adults.

The preventive health questions are asked in alternating years and will not be included in this file; they will be included in the 2022 SAQ. In 2021, questions regarding quality of health care, general health questions, and questions about health-related attitudes were asked in the SAQ and are included in this file.

The 2021 SAQ was fielded during Panel 23 Round 8, Panel 24 Round 6, Panel 25 Round 4, and Panel 26 Round 2 of the 2021 MEPS data collection.

All adults age 18 and older as of the Round 2, 4, 6, or 8 interview date (AGE42X >= 18) in MEPS households were asked to complete an SAQ. The questionnaires were administered in late 2021 and early 2022.

The variable SAQELIG indicates the person’s eligibility status for the SAQ. SAQELIG was used to construct the variables based on the SAQ data. SAQELIG was coded “0” (Not Eligible for

SAQ) if there was no record for the person in the round, if the person was deceased or institutionalized, moved out of the U.S., moved to a military facility, if the person's disposition status was inapplicable, or if the person was less than 18 years old. SAQELIG was coded "1" (Eligible for SAQ and Has SAQ Data) if an SAQ record existed for the person in Round 2 (for Panel 26), Round 4 (for Panel 25), Round 6 (for Panel 24), or Round 8 (for Panel 23). SAQELIG was coded "2" (Eligible for SAQ, but No SAQ Data) if no SAQ record existed for the person in the round. This variable was used as a building block for all other constructed SAQ variables.

A question on the form asked if the respondent was the person represented in the form. If a person was unable to respond to the SAQ, the questionnaire was completed by a proxy. The relationship of the proxy to the adult represented in the questionnaire is indicated by the variable ADPROX42. Prior to 2015, the variable ADPRX42 indicated the relationship of the proxy to the adult. Starting in 2015, the response categories for proxy relationship were collapsed in a new variable ADPRXY42. ADPROX42 was coded "1" (Self-administered) if the respondent was the person represented in the questionnaire. A code of "-1" (Inapplicable) was assigned if a person was not eligible or was eligible but no data existed (SAQELIG = 0 or 2).

If a person was not assigned a positive SAQ weight, all SAQ variables except SAQELIG were coded "Inapplicable" (-1). When a gate question answer was set to "No" (2), follow-up variables based on the gate question were coded as "Inapplicable" (-1). When a gate question answer was set to "Refused" (-7) or "Don't Know" (-8), follow-up variable answers were left as reported. A special weight variable (SAQWT21F) has been designed to be used with the SAQ for persons who were age 18 and older at the interview date. This weight adjusts for non-response and weights to the U.S. civilian noninstitutionalized population (see Section 3.0 "Survey Sample Information" for details). The variables created from the SAQ begin with "AD," again excepting SAQELIG.

Health Care Quality (included in alternating years only)

The health care quality measures in the SAQ were taken from the health plan version of CAHPS®, an AHRQ-sponsored family of survey instruments designed to measure quality of care from the consumer's perspective. All of the variables refer to events experienced in the last 12 months and were asked of adults age 18 and older. The variables included from the CAHPS® are:

- ADILCR42 - Had an illness, injury or condition needing care right away from a clinic, emergency room or doctor's office
- ADILWW42 - If ADILCR42 = 1, how often got care right away
- ADRTCR42 - Any appointment was made with a doctor or clinic for health care
- ADRTWW42 - If ADRTCR42 = 1, how often got an appointment for health care as soon as they thought it was needed
- ADAPPT42 - Number of times went to doctor's office or clinic to get care
- ADLIST42 - If ADAPPT42 > 0, how often health providers listened carefully to you

- ADEXPL42 - If ADAPPT42 > 0, how often health providers explained things in a way that was easy to understand
- ADRESP42 - If ADAPPT42 > 0, how often providers showed respect for what you had to say
- ADPRTM42 - If ADAPPT42 > 0, how often health providers spent enough time with you
- ADINST42 - If ADAPPT42 > 0, whether doctors or other health providers gave instructions about what to do about a specific illness or health condition
- ADEZUN42 - If ADINST42 = 1, how often the advice given by doctors or other health providers was easy to understand
- ADTLHW42 - If ADINST42 = 1, how often doctors or other health providers asked you to describe how you are going to follow their instructions
- ADFFRM42 - If ADAPPT42 > 0, whether had to fill out or sign any forms at the doctor's office or other health provider's office
- ADFHLP42 - If ADFFRM42 = 1, how often you were offered help with filling out forms at the office
- ADHECR42 - If ADAPPT42 > 0, rating of healthcare from all doctors and other health providers, from 0 (worst health care possible) to 10 (best health care possible)
- ADSPCL42 - Needed to see a specialist
- ADSNP42 - If ADSPCL42 = 1, how easy to see a specialist

General Health (included in alternating years only)

- ADSMOK42 - Currently smoke
- ADNSMK42 - If ADSMOK42 = 1, doctor advised you to quit smoking

Health Status

The SAQ contained three measures of health status: the Veteran RAND (VR-12), a registered trademark, the Kessler Index (K6) of non-specific psychological distress, and the Patient Health Questionnaire (PHQ-2). More information about the VR-12 is available through the [Boston University School of Public Health website](#). Key references for these three measures are Kessler et al. (2002), Kroenke et al. (2003), Selim et al. (2018) and Selim et al. (2009).

Veterans RAND 12 Version (VR-12)

The Veterans RAND 12 Item Health Survey (VR-12[©]) is a self-administered health survey comprising 12 items used to measure health related quality of life, to estimate disease burden and to evaluate disease-specific impact on general and selected populations.

The VR instrument uses five-point ordinal response choices for four items in the VR-12[®]. Response choices are five-point response choices: “no, none of the time”, “yes, a little of the time”, “yes, some of the time”, “yes, most of the time” and “yes, all of the time.” These answers then contribute to the scales for role limitations due to physical and emotional problems (PCS) and the physical and mental summary scores (MCS).

In analyzing data from the VR-12, the standard approach is to form two summary scores based on responses to the 12 questions. The standard scoring algorithms for both the Physical Component Summary (PCS) and the Mental Component Summary (MCS) incorporate information from all 12 questions. However, the PCS weights more heavily responses to the following questions: ADGENH42, ADDAYA42, ADCLIM42, ADACLS42, ADWKLM42, and ADPAIN42. The MCS weights more heavily responses to the following questions: ADPRST42, ADPCFL42, ADEMLS42, ADMWDF42, and ADSOCA42. The computer programs to create VR scales and PCS/MCS summaries are copyrighted (all rights reserved) by the Trustees of Boston University to ensure the integrity of the assessments.

The comparability of the 2017 MEPS VR-12 PCS and MCS summary scores from the standard scoring algorithm and the SF-12v2 PCS and MCS summary scores obtained from prior years of MEPS was assessed, and it was determined that the scores were misaligned. A bridging algorithm specific to MEPS was developed by a team at the Boston University School of Public Health. The goal of this bridging algorithm was to align the VR-12 PCS and MCS scores from the 2017 MEPS as closely as possible with the SF-12v2 PCS and MCS scores from prior MEPS years across a wide range of MEPS subpopulations. This bridging algorithm was applied to the VR-12 PCS and MCS score variables (VPCS42 and VMCS42) available on this data file.

The PCS and MCS cannot be computed directly if a person has missing data for any of the twelve items. A proprietary method was used for imputing the PCS and MCS scores if some data are missing. The bridging algorithm used for these measures was developed to be tolerant of missing data in item responses when computing PCS and MCS scores.

Therefore, the variables VPCS42 and VMCS42 include some cases in which the scores have been imputed. Some cases were unable to be scored in the bridging algorithm due to the amount of missing data in item responses; these cases have VRFLAG42 = 0 (No). VRFLAG42 indicates whether the physical component summary, VPCS42, or the mental component, VMCS42, was imputed for a respondent. Persons who were not eligible for the SAQ, or who were eligible but for whom no data existed based on SAQELIG, or who did not have a positive SAQ weight, were set to “Inapplicable” (-1) for VRFLAG42, VPCS42 and VMCS42. Any remaining persons who could not be scored were set to “Cannot be Computed” (-15) for VPCS42 and VMCS42. Additionally, beginning in 2017, there are no negative score values of VPCS42 and VMCS42 because they are both top- and bottom-coded.

More information on the VR-12 can be found on the [Boston University website VR-12 page](#). The report containing information on the methodology used for the bridging algorithm can be requested from mepsprojectdirector@ahrq.hhs.gov.

The VR-12 questions are as follows:

- ADGENH42 - General health today
- ADDAYA42 - During a typical day, limitations in moderate activities
- ADCLIM42 - During a typical day, limitations in climbing several flights of stairs
- ADACLS42 - During past 4 weeks, as result of physical health, accomplished less than would like
- ADWKLM42 - During past 4 weeks, as result of physical health, limited in kind of work or other activities
- ADEMLS42 - During past 4 weeks, as result of emotional problems, accomplished less than you would like
- ADMWDF42 - During past 4 weeks, as result of emotional problems, did work or other activities less carefully than usual
- ADPAIN42 - During past 4 weeks, pain interfered with normal work outside the home and housework
- ADPCFL42 - During the past 4 weeks, felt calm and peaceful
- ADENGY42 - During the past 4 weeks, had a lot of energy
- ADPRST42 - During the past 4 weeks, felt downhearted and blue
- ADSOCA42 - During the past 4 weeks, physical health or emotional problems interfered with social activities

Non-Specific Psychological Distress

The 2021 SAQ includes six mental health-related questions, using the “K-6” scale developed by R.C. Kessler and colleagues. These questions assess the person’s non-specific psychological distress during the past 30 days.

The non-specific psychological distress variables are as follows:

- ADNERV42 - During the past 30 days, how often felt nervous
- ADHOPE42 - During the past 30 days, how often felt hopeless
- ADREST42 - During the past 30 days, how often felt restless or fidgety
- ADSAD42 - During the past 30 days, how often felt so sad that nothing could cheer the person up
- ADEFRT42 - During the past 30 days, how often felt that everything was an effort

ADWRTH42 - During the past 30 days, how often felt worthless

Kessler Index (K6)

A summary of the six variables above provides an index to measure non-specific, rather than disorder-specific, psychological distress, using the following values:

- 0 None of the Time
- 1 A Little of the Time
- 2 Some of the Time
- 3 Most of the Time
- 4 All of the Time

The index, called K6SUM42, is a summation of the values of the six variables above. The higher the value of K6SUM42, the greater the person's tendency towards mental disability.

Patient Health Questionnaire (PHQ-2)

The 2021 SAQ includes two additional mental health questions. These questions assess the frequency of the person's depressed mood and decreased interest in usual activities.

ADINTR42 - During the past two weeks, bothered by having little interest or pleasure in doing things

ADDPRS42 - During the past two weeks, bothered by feeling down, depressed, or hopeless

PHQ242 is a summation of the values of the two variables above, with scores ranging from 0 through 6. The higher the value of PHQ242, the greater the person's tendency towards depression. Kroenke et al. (2003) suggest a score of 3 as the optimal cut point for screening purposes. Note that these items are intended as a screening measure for depression and are not equivalent to a DSM-V diagnosis of depression.

The language in which the SAQ was completed is indicated by the variable ADLANG42. If the English version of the SAQ was completed, ADLANG42 was coded "1" (English Version SAQ Was Administered). If the Spanish version of the SAQ was completed, or if the English version was translated into Spanish, ADLANG42 was coded "2" (Spanish Version SAQ Was Administered). If the language in which the SAQ was administered could not be determined from the data, ADLANG42 was coded "-15" (Cannot be Computed).

The month and year the SAQ was completed are indicated by the variables ADCMPM42 and ADCMPY42, respectively.

When using the SAQ variables in analysis, weights specific to these questions should be used (SAQWT21F). For persons who are not assigned a positive SAQ weight, the SAQ variables are recoded to "Inapplicable" (-1). Please see Section 3.0 "Survey Sample Information" for details.

Diabetes Care Survey (DCS)

The Diabetes Care Survey (DCS) is a self-administered paper-and-pencil questionnaire fielded during Panel 23 Round 9, Panel 24 Round 7, Panel 25 Round 5, and Panel 26 Round 3. Households received a DCS based on their response to DIABDX_M18 in the Priority Conditions Enumeration (PE) section of the CAPI instrument, which asks whether the person was ever told by a doctor or health professional that they had diabetes. Note that only those 18 years or older were asked to complete a DCS questionnaire.

The DCS asks the same question as DIABDX_M18 with responses summarized in the variable DSDIA53. DSDIA53 confirms that the person has ever been told by a health professional that they had diabetes or sugar diabetes. Every year, a small number (50) of people answer no to the diabetes diagnosis question (DCS.DIABDIAG) on the DCS. These people have DSDIA53 initially set to “2” (No). DCS.DIABDIAG is used in the development of the diabetes weight (DIABWyyF); if the person has DCS.DIABDIAG = “2” (No) they do not receive a diabetes weight (DIABWyyF = 0). In the final stage of DCS variable construction, DCS constructed variables, excluding the eligibility variable (DCSELIG) were recoded to “-1” where DIABWyyF = 0. For these cases, DIABDX_M18 = YES (1) but DSDIA53 = NO (2). The DCS data are unedited, and, therefore, these and other data inconsistencies remain in the data. For all persons 17 years of age or younger, all the DCS variables are set to “-1” (Inapplicable) because there is not an appropriate weight included on the file to make national estimates for this population.

DSA1C53 indicates the number of times the respondent reported having a hemoglobin A1c blood test in 2021. Note that, prior to 2005, DSA1C53 did not reflect whether the person had a hemoglobin A1c blood test, only whether the person had a hemoglobin A1c test. DSFT2253, DSFT2153, DSFT2053, DSFB2053, and DSFTNV53 indicate whether the respondent reported having their feet checked for sores or irritations: in 2022, in 2021, in 2020, before 2020, or never, respectively. DSEY2253, DSEY2153, DSEY2053, DSEB2053 and DSEYNV53 indicate whether the respondent reported having an eye exam in which the pupils were dilated: in 2022, in 2021, in 2020, before 2020, or never, respectively. DSCH2253, DSCH2153, DSCH2053, DSCB2053, and DSCHNV53 indicate the last time the respondent reported having their blood cholesterol checked: in 2022, in 2021, in 2020, before 2020, or never, respectively. DSFL2253, DSFL2153, DSFL2053, DSVB2053, and DSFLNV53 indicate when the person got a flu vaccination including the flu vaccine nasal spray: in 2022, in 2021, in 2020, before 2020, or never, respectively. DSKIDN53 and DSEYPR53 ascertain whether the diabetes has caused kidney or eye problems, respectively. DSDIET53, DSMED53, and DSINSU53 indicate if the respondent reported being treated for his/her diabetes by the following methods: diet, oral medications, or insulin, respectively.

The five variables that assess different ways the person with diabetes can learn about diabetes care are: DSCPCP53 (learned care from a primary care provider), DSCNPC53 (learned care from a provider not in the person’s primary care practice), DSCPHN53 (learned care from a phone call with a provider), DSCINT53 (learned care from reading about it on the internet), and DSCGRP53 (learned care by taking a group class). Creation of these variables is based on the answer to a gate question, which asks, “During the last 12 months, have you learned how to take care of your diabetes?” Please note that there is no variable listed in the codebook to indicate the answer to that question, since it is only used for creation of the follow-up variables DSCPCP53,

DSCNPC53, DSCPHN53, DSCINT53, and DSCGRP53. These follow-up variables are set to Inapplicable (-1) for persons who report *not* having learned how to take care of their diabetes during the last 12 months. The variable DSCONF53 indicates how confident the person is in treating their diabetes. Those variables that indicate a range of care outside the data year may represent persons with additional information included on the 2020 or the 2022 Full Year Consolidated PUF. Additional data for the second-year panel may be available on the 2020 PUF.

If a person was unable to respond to the DCS, the questionnaire was completed by a proxy (DSPRX53 = 1). A special weight variable (DIABW21F) has been designed to be used with DCS data. This weight adjusts for DCS nonresponse and weights to the number of diabetics in the U.S. civilian noninstitutionalized population in 2021 (see Section 3.0 “Survey Sample Information” for details). When using these variables in analysis, weights specific to each of these sets of questions should be used (SAQWT21F and DIABW21F). For persons who are not assigned a positive DCS weight, the DCS variables are recoded to “Inapplicable” (-1). Please see Section 3.0 “Survey Sample Information” for details.

Social Determinants of Health Survey (SDOH)

The Social Determinants of Health Survey (SDOH) includes questions about the social determinants of health, such as housing, financial well-being, food security, social support, discrimination, and physical and sexual violence. The SDOH constitutes the first MEPS instrument to be administered both as a paper-and-pencil and web questionnaire.

The SDOH was fielded during Panel 23 Round 7, Panel 24 Round 5, Panel 25 Round 3, and Panel 26 Round 1 of the MEPS data collection.

All adults age 18 and older as of the Round 1, 3, 5, or 7 interview date (AGE31X >= 18) in MEPS households were asked to complete the SDOH. No gate questions were included, all questions were asked of all SDOH respondents, regardless of age, sex, or health status.

The variable SDOHELIG indicates the person’s eligibility status for the SDOH. SDOHELIG was used to construct the variables based on the SDOH data. SDOHELIG was coded “0” (Not Eligible for SDOH) if there was no record for the person in the round, if the person was not key, if the person was deceased or institutionalized, moved out of the U.S., moved to a military facility, if the person’s disposition status was inapplicable, or if the person was less than 18 years old. SDOHELIG was coded “1” (Eligible for SDOH and Has SDOH Data) if an SDOH record existed for the person in Round 1 (for Panel 26), Round 3 (for Panel 25), Round 5 (for Panel 24), or Round 7 (for Panel 23). SDOHELIG was coded “2” (Eligible for SDOH, but No SDOH Data) if no SDOH record existed for the person in the round. This variable was used as a building block for all other constructed SDOH variables.

A question that was only included on the paper form asked if the respondent was the person represented in the form. If a person was unable to respond to the SDOH, the paper questionnaire was completed by a proxy. The relationship of the proxy to the adult represented in the questionnaire is indicated by the variable SDPROX42. SDPROX42 was coded “1” (Self-administered) if the respondent was the person represented in the questionnaire. A code of “-1” (Inapplicable) was assigned if a person was not eligible or was eligible but no data existed

(SDOHELIG = 0 or 2). Month and year of survey completion are included - SDCMPPM and SDCMPY.

A special weight variable (SDOHWT21F) has been designed to be used with the SDOH for persons who were age 18 and older at the interview date. This weight adjusts for non-response and weights to the U.S. civilian noninstitutionalized population (see Section 3.0 “Survey Sample Information” for details). The variables created from the SDOH begin with “SD.” If a person was not assigned a positive SDOH weight, all SDOH variables except SDOHELIG were coded “Inapplicable” (-1). See [Social and Health Experiences Sources](#) on the MEPS website for a list of sources used for question wording.

SDLIFE -	How satisfied with life
SDHOME -	How satisfied with house/apartment where person lives
SDMEDCARE -	Rate neighborhood availability of places to get medical care
SDPARKS -	Rate neighborhood availability of parks and playgrounds
SDHLTHFOOD -	Rate neighborhood availability of places to buy healthy food
SDSFCRIME -	Rate neighborhood safety from crime and violence
SDPUBTRANS -	Rate neighborhood access to public transportation
SDAFRDHOME -	Rate neighborhood availability of affordable housing
SDNOTRANS -	During last 12 months, has lack of transportation kept person from medical appointments, meetings, work, or getting things needed for daily living
SDLATERENT -	During last 12 months, household’s rent/mortgage was late or unpaid because household could not afford to pay
SDLATEUTIL -	During last 12 months, household unable to pay electric, gas, oil, or water bills on time
SDSHUTUTIL -	During last 12 months, home utilities were threatened to be shut off
SDPROBPEST -	The place where person lives has problems with pests
SDPROBMOLD -	The place where person lives has problems with mold
SDPROBLEAD -	The place where person lives has problems with lead paint or pipes
SDPROBHEAT -	The place where person lives has problems with lack of heat
SDPROBCOOK -	The place where person lives has problems with oven or stove not working

SDPROBSMKDET -	The place where person lives has problems with smoke detectors missing or not working
SDPROBLEAKS -	The place where person lives has problems with water leaks
SDPROBNONE -	The place where person lives has problems with none of the above
SDWRRYFD -	During last 12 months, worried about food running out before person got money to buy more
SDNOFOOD -	During last 12 months, food bought didn't last and person did not have money to get more
SDPAYBASICS -	How hard is it to pay for basics like food, housing, medical care, and heating
SDUNEXPEXP -	How confident able to come up with \$400 if an unexpected expense arose within the next month
SDMISSCCLN -	During last 12 months, missed a payment on a credit card or loan (not including mortgage)
SDDEBT -	During last 12 months, contacted by debt collection agency
SDDAYEXER -	During last 30 days, how many days per week engaged in moderate exercise
SDMINSEXER -	Average amount of minutes spent doing moderate exercise
SDSTRESS -	Rate current stress level
SDFAMILY -	How much help expected to receive from family if problem arose (for example, sickness or moving)
SDFRIENDS -	How much help expected to receive from friends if problem arose (for example, sickness or moving)
SDCOMM -	How much help expected to receive from community if problem arose (for example, sickness or moving)
SDTLKPHN -	How often talk on the phone (or by video) to family, friends, or neighbors in a typical week
SDGETTGT -	How often get together with friends or relatives in a typical week
SDCHURCH -	How often attend church or religious services (in-person or online)
SDCLUBORG -	How often attend meetings of clubs or organizations (in-person or online)

SDCOMPAN -	How often feel the lack of companionship
SDLEFTOUT -	How often feel left out
SDISOL -	How often feel isolated from others
SDENICPROD -	Ever used an electric nicotine product
SDDSCRMDR -	Ever experienced discrimination at doctor's office, clinic, or hospital
SDDSCRMWK -	Ever experienced discrimination at work
SDDSCRMJOB -	Ever experienced discrimination when applying for jobs
SDDSCRMHHS -	Ever experienced discrimination when trying to rent or buy housing
SDDSCRMPOL -	Ever experienced discrimination when interacting with law enforcement
SDDSCRMPUB -	Ever experienced discrimination when applying for social services or public assistance
SDDSCRMSTR -	Ever experienced discrimination at a restaurant or store
SDPHYSHURT -	How often does anyone physically hurt person
SDINSULT -	How often does anyone insult or talk down to person
SDTHRHARM -	How often does anyone threaten person with harm
SDSCREAM -	How often does anyone scream or curse at person
SDHMDEPR -	When under 18 years of age, lived with someone who was depressed, mentally ill, or suicidal
SDHMALC -	When under 18 years of age, lived with a problem drinker or alcoholic
SDHMDRG -	When under 18 years of age, lived with someone who used illegal street drugs or abused prescription medication
SDHMJAIL -	When under 18 years of age, lived with someone who served time in prison, jail, or correctional facility
SDHMDIV -	Parents separated or divorced
SDHMBEAT -	How often did parents or adults in home slap, hit, kick, punch, or beat each other up

SDHURTCHLD -	How often did parents or adults in home hit, beat, kick, or physically hurt person
SDINSCHLD -	How often did parents or adults in home swear, insult, or put person down
SDTCHCHLD -	How often did an adult or someone at least 5 years older touch person sexually
SDTCHADLT -	How often did an adult or someone at least 5 years older try to make person touch them sexually
SDFRCSXCH -	How often did an adult or someone at least 5 years older force the person to have sex

2.5.7 Disability Days Indicator Variables (DDNWRK21-OTHNDD21)

The Disability Days (DD) section of the core interview contains questions about time lost from work because of a physical illness or injury, or a mental or emotional problem. Data were collected on each individual in the household. These questions were repeated in each round of interviews; this file contains data from Rounds 7, 8, and 9 of the MEPS Panel 23, initiated in 2018; Rounds 5, 6, and 7 of the MEPS Panel 24, initiated in 2019; Rounds 3, 4, and 5 of the MEPS Panel 25, initiated in 2020; and Rounds 1, 2, and 3 of the MEPS Panel 26, initiated in 2021.

Beginning in FY 2015, annualized versions of these variables are constructed for release rather than the previously released versions, which were round- and panel-specific. The number at the end of the variable name (21) identifies the variable as representing data from 2021. Due to confidentiality concerns, the annual Disability Days variables, which represent the number of days a person missed work (DDNWRK21 and OTHNDD21), are top-coded to mask values that exceed the top one-half of one percent of the population.

These annual variables use building block variables for construction, which represent an individual panel within the data year. The reference period for the Disability Days questions is the time period between the beginning of the panel or the previous interview date and the current interview date. Analysts should be aware that Round 7 of Panel 23, Round 5 of Panel 24, and Round 3 of Panel 25 are conducted across years. The Disability Days variables reflect only the data pertinent to the calendar year (i.e., the current delivery year of 2021). Analysts who are interested in examining Disability Days data across years can link to other person-level PUFs using the DUPERSID.

The flow of the Disability Days section relies on the person's age as of the interview date. Therefore, the round-specific constructed age variables (AGE31X, AGE42X, and AGE53X) are used to construct the comparable round-specific Disability Days building block variables. Due to the age-specific nature of the Disability Days section, age data from other rounds are not used should the person's age for the round be missing.

The variable DDNWRK21 represents the number of times the person lost a half-day or more from work because of illness, injury, or mental or emotional problems during the calendar year. A response of “no work days lost” was coded zero; if the person did not work, this variable was coded -1 (Inapplicable). The analyst should note that there are cases where EMPST## = 1 or 2 (has current job or job to return to) where DDNWRK21 contains a positive value, indicating the number of times the person lost a half-day or more from work. This is because the responses to the Disability Days questions are independent of the responses to the employment questions. Persons who were less than 16 years old or whose age is missing (AGE##X is set to -1) were not asked about work days lost, thus this variable is coded -1 (Inapplicable) for these persons.

A final set of variables indicates if an individual took a half-day or more off from work to care for the health problems of another individual in the family and the number of days missed. OTHDYS21 indicates if a person missed work because of someone else’s illness, injury, or health care needs, for example to take care of a sick child or relative. This variable has three possible answers: yes - missed work to care for another (coded 1); no - did not miss work to care for another (coded 2); or the person does not work (coded -1), based on the setting of DDNWRK21. Persons younger than 16 and persons whose age is missing were not asked this question and are also coded as -1 (in a small number of cases this was not done for the 1996 data, the analyst will need to make this edit when doing longitudinal analyses).

OTHNDD21 indicates the number of days in which work was lost because of another’s health problem. Persons younger than 16, those whose age is missing, those who do not work, and those who answer “no” to OTHDYS21 are skipped out of OTHNDD21 and receive a code of -1.

Note that, because Disability Days variables use only those Round 7 of Panel 23, Round 5 of Panel 24, and Round 3 of Panel 25 data pertinent to the data year, it is possible to have a person report missing work to care for the health problems of another individual (OTHDYS21 = 1) but report no days missed (OTHNDD21 = 0). This combination indicates that the person did not miss those workdays during the data year. Editing was done on these variables to preserve the skip patterns. No imputation was done for those with missing data.

The variables DDNWRK21 and OTHNDD21 are annualized variables derived from responses to questions at each interview round. If the round-specific responses included a mix of missing values (-1, -7, -8, or 0), then the annualized variables were set to -15 CANNOT BE COMPUTED on the PUF. This editing results in elevated rates of -15 values for these variables compared with other variables on the PUF.

2.5.8 Access to Care Variables (ACCELI42- BOOSTERSHOT53)

The variables ACCELI42 through AFRDPM42 describe data from the Access to Care (AC) section of the MEPS HC questionnaire, which was administered in Panel 23 Round 8, Panel 24 Round 6, Panel 25 Round 4, and Panel 26 Round 2. This supplement gathers information on family members’ usual source of health care; characteristics of usual source of health care providers; access to and satisfaction with the usual source of health care provider; and affordability of medical treatment, dental treatment, and prescription medicines.

The variable ACCELI42 indicates whether persons were eligible to receive the Access to Care questions. Persons with ACCELI42 set to '-1' (Inapplicable) should be excluded from estimates made with the Access to Care data.

The COVID (CV) section of the MEPS HC questionnaire was administered in Panel 23 Rounds 7 - 9, Panel 24 Rounds 5 - 7, Panel 25 Rounds 3 - 5, and Panel 26 Rounds 1 - 3. This supplement gathers information on whether a person was delayed in getting medical care (CVDLAYCA31, CVDLAYCA42, CVDLAYCA53), dental treatment (CVDLAYDN31, CVDLAYDN42, CVDLAYDN53), and prescription medicine (CVDLAYPM31, CVDLAYPM42, CVDLAYPM53) due to the coronavirus pandemic, as well as whether the person ever acquired a COVID-19 vaccination (CVVACCINE42, CVVACCINE53) and ever had a COVID-19 booster vaccination (BOOSTERSHOT53).

Family Members' Usual Source of Health Care

For each individual family member, the AC section ascertains whether there is a particular doctor's office, clinic, health center, or other place that the individual usually goes to if they are sick or need advice about their health (HAVEUS42).

PRACTP42 indicates whether a usual source of care (USC) provider has their own practice that is not part of a group practice, health center, clinic, or other facility. For those family members who have a USC provider, AC30 ascertains the type of practice:

- 1 Own Practice, Not Part of Group/Facility
- 2 Practice Associated with Group/Facility

YNOUSC42_M18 indicates the main reason why a person does not have a USC provider. For those family members who do not have a USC provider, question AC40 ascertains the main reason why:

- 1 Seldom or Never Sick
- 2 Recently Moved to Area
- 3 Just Changed Insurance Plans
- 4 No Health Insurance, Oth Insurance-Related Issue
- 5 Don't Know Where to Go for Care
- 6 USC in This Area No Longer Available
- 7 Likes to Go to Different Places for Different Health Needs
- 8 Don't Use Doctors/Treat Self
- 9 Cost of Medical Care

- 10 No Health Insurance
- 91 Other Reason

In 2018, YNOUSC42 was renamed to YNOUSC42_M18 because the list of answer categories changed.

Characteristics of Usual Source of Health Care Providers

The AC section collects information about the different characteristics of each unique USC provider for a given family. If a person does not have a USC provider (HAVEUS42 is set to ‘2’ (No), ‘-7’ (Refused), ‘-8’ (Don’t Know) or ‘-15’ (Cannot be Computed)), then these variables are set to ‘-1’ (Inapplicable).

The basis for the AC provider questions is PROVTY42_M18. This variable indicates whether the person’s provider is a facility (‘1’), a person (‘2’), or a person-in-facility (‘3’). PROVTY42_M18 is a copy of PROVTYPE_M18 (Provider Type) for persons who have a USC provider. Depending on how PROVTYPE_M18 is set, persons are asked about the provider’s location, the provider’s personal characteristics (e.g., race), the provider’s accessibility, and the person’s satisfaction with the provider. In 2018, PROVTY42 was renamed PROVTY42_M18 because of changes to CAPI.

Provider Location

Two variables indicate the location of the provider. For facility or person-in-facility type providers, PLCTYP42 indicates whether the person’s facility is a Hospital Clinic or Outpatient Department (‘1’), Hospital Emergency Room (‘2’), or Other Kind of Place (‘3’). According to CAPI flow, persons do not report the type of facility for person-type providers; therefore, if PROVTY42_M18 is set to ‘2’ (Person), PLCTYP42 is set to ‘-1’ (Inapplicable).

For all provider types, including person-type, LOCATN42 indicates whether the person’s provider is located in an Office (‘1’), a Hospital but Not the Emergency Room (‘2’), or a Hospital Emergency Room (‘3’).

Personal Characteristics of Providers

For person and person-in-facility type providers, TYPEPE42 indicates what type of doctor or other medical provider the person’s provider is. The possible values include:

- 1 MD - General/Family Practice
- 2 MD - Internal Medicine
- 3 MD - Pediatrics
- 4 MD - OB/Gyn
- 5 MD - Surgery
- 6 MD - Other

- 7 Chiropractor
- 8 Nurse
- 9 Nurse Practitioner
- 10 Physician's Assistant
- 11 Other Non-MD Provider
- 12 Unknown
- 13 MD - Cardiologist
- 14 Doctor of Osteopathy
- 15 MD - Endocrinologist
- 16 MD - Gastroenterologist
- 17 MD - Geriatrician
- 18 MD - Nephrologist
- 19 MD - Oncologist
- 20 MD - Pulmonologist
- 21 MD - Rheumatologist
- 22 Psychiatrist / Psychologist
- 23 MD - Neurologist
- 24 Alternative Care Provider

TYPEPE42 is constructed using variables collected at several questions: AC70 “Is provider a medical doctor?” (PROV.MEDTYPE_M18); AC80 “Is provider a nurse, nurse practitioner, physician’s assistant, midwife, or some other kind of person?” (PROV.OTHTYPE_M18); and AC90 “What is provider’s specialty?” (PROV.MDSPECLT_M18). If respondents choose ‘91’ (Other) at AC80 or AC90, they are asked at AC80OS or AC90OS, respectively, to provide a verbal explanation of the type of provider or medical doctor. These text strings can be recoded to one of the existing categorical values listed above or, if the frequency of the response warrants it, additional categorical values. Recoding is described in greater detail below.

The AC section also collects demographic information about person and person-in-facility type providers (PROVTY42 = 2 or 3). Six variables indicate the provider’s race: WHITPR42 (White), BLCKPR42 (Black/African American), ASIANP42 (Asian), NATAMP42 (Indian/Native American/Alaska Native), PACISP42 (Other Pacific Islander) and OTHRCP42 (Other Race).

The respondent may choose more than one race for a single provider. These variables reflect the answer categories given at AC110.

In addition to the race variables, two other demographic variables are created: HSPLAP42 indicates whether or not the provider is Hispanic or Latino, and GENDRP42 indicates whether the provider is Male ('1') or Female ('2').

Using Constructed Variables to Describe the Usual Source of Care Provider

These variables describing a person's USC provider can be used in combination to present a broader picture of the provider. For example, a person-in-facility provider with a particular person named who is a white, Hispanic, female pediatrician, with no other race specified, and whose location is in a hospital is coded as:

PROVTY42_M18 = 3	ASIANP42 = 2
PLCTYP42 = 1	NATAMP42 = 2
TYPEPE42 = 3	PACISP42 = 2
HSPLAP42 = 1	OTHRCP42 = 2
WHITPR42 = 1	GENDRP42 = 2
BLCKPR42 = 2	LOCATN42 = 2

Access to and Satisfaction with the Provider

The AC section collects information regarding the person's ability to access the USC provider as well as the person's satisfaction with the USC provider.

Access to the Provider

TMTKUS42 indicates how long it takes the person to travel to the USC provider: Less Than 15 Minutes ('1'), 15 to 30 Minutes ('2'), 31 to 60 Minutes ('3'), 61 to 90 Minutes ('4'), 91 Minutes to 120 Minutes ('5'), or More than 120 Minutes ('6').

OFFHOU42, PHNREG42, and AFTHOU42 assess aspects of the provider that may make it difficult for the person to get in contact with the USC provider. OFFHOU42 indicates whether the provider has office hours at night or on the weekend. The remaining two variables reflect the person's rating of the difficulty of accessing the USC provider by phone (PHNREG42), and after hours (AFTHOU42). The person has the following choices: Very Difficult ('1'), Somewhat Difficult ('2'), Not Too Difficult ('3'), or Not at All Difficult ('4').

Satisfaction with the Provider

These variables reflect the person's satisfaction with the USC provider. The person's level of satisfaction with the USC provider is examined in four ways: Does the USC provider: usually ask about prescription medications and treatments other doctors may give them (TREATM42), ask the person to help make decisions between a choice of treatments (DECIDE42), present and

explain all options to the person (EXPLOP42), and speak the person's language or provide translator services (PRVSPK42).

PRVSPK42 is set to a value other than '-1' (Inapplicable) for persons eligible for the Access to Care supplement who had a usual source of care, were identified as speaking a language other than English at home (OTHLGSPK = '1') and speaking English either "Not Well" or "Not at All" (HWELLSPK = '3' or '4'). PRVSPK42 is set to '-1' (Inapplicable) for all persons not meeting these criteria or who were deceased, institutionalized, or younger than 5 years of age.

If the person was under 5 years old in Round 1 and age 5 in Round 2 of the first year panel or Round 4 of the second year panel, and the source data are missing, PRVSPK42 was set to '-1' (Inapplicable); if the source data are available, PRVSPK42 was set per specifications.

Affordability of Medical Care, Dental Care, and Prescription Medicines

The Access to Care supplement gathers information on whether care was not received or was delayed because of cost in the past 12 months. These questions are split into three sections inquiring about medical care, dental care, and prescription medicines. Each section inquires whether the person did not receive care because they could not afford it (AFRDCA42, AFRDDN42, AFRDPM42). The affordability variables indicate with a value of '1' (Yes) that the person needed care but was unable to afford it, a value of '2' (No) that the person did not have any unmet needs for that type of care because of the cost.

Respondents were also asked if anyone in the household delayed receiving care because of worry about cost (DLAYCA42, DLAYDN42, DLAYPM42). The delay variables indicate with a value of '1' (Yes) that the person was delayed in receiving that type of care because of worry about the cost, and a value of '2' (No) for these variables indicates that the person was not delayed in seeking that type of care because of the worry about the cost.

Editing the Access to Care Variables

Editing consisted primarily of logical editing for consistency with skip patterns. Other editing included the construction of new response values and new variables describing the recoding of "other specify" text items into existing or new categorical values, which are described below.

Not all variables or categories that appear in the Access to Care section of the HC questionnaire are included on the file, as some small cell sizes have been suppressed to maintain confidentiality.

Recoding of Additional Other Specify Text Items

For Access to Care items AC80 and AC90, the "other specify" text responses were reviewed and coded as an existing or new value for the related categorical variable (AC80 and AC90).

OTHTYPE_M18 and MDSPECLT_M18 are used to construct the variable TYPEPE42. These variables' text strings can be recoded to each other's categories. For example, for persons who indicate that their USC provider is not a medical doctor (PROV.MEDTYPE = '2'), the other type of USC provider is other (PROV.OTHTYPE = '91'), and the text string collected is

“GYNECOLOGIST,” TYPEPE42 would be set to ‘4’ (MD - OB/GYN) instead of ‘11’ (OTHER NON-MD PROVIDER.)

Delayed Medical Care, Dental Care, and Prescription Medicines due to the Coronavirus Pandemic

The CV section was administered in Panel 23 Rounds 7 - 9, Panel 24 Rounds 5 - 7, Panel 25 Rounds 3 - 5, and Panel 26 Rounds 1 - 3. The CV section gathered information on family members’ abilities to receive treatment without delay from March 2020 (for Panel 23 Round 7, Panel 24 Round 5, Panel 25 Round 3, and Panel 26 Round 1) and the last interview date for all other panels. The following table summarizes the panels and rounds used in construction of CVDLAYCArr, CVDLAYDNrr, and CVDLAYPMrr variables, where rr indicates a combination of rounds as described in the table below.

rr	Panel 23	Panel 24	Panel 25	Panel 26
31	Round 7	Round 5	Round 3	Round 1
42	Round 8	Round 6	Round 4	Round 2
53	Round 9	Round 7	Round 5	Round 3

These questions asked whether anyone in the household delayed receiving care because of the coronavirus pandemic. If the respondent answered ‘1’ (Yes), they were asked to identify who in the household delayed care. Within a household that had care delayed, the variables CVDLAYCA31, CVDLAYCA42, CVDLAYCA53 (Delay Med Care For COVID R3/1, R4/2, R5/3), CVDLAYDN31, CVDLAYDN42, CVDLAYDN53 (Delay Getting Dental For COVID R3/1, R4/2, R5/3), and CVDLAYPM31, CVDLAYPM42, CVDLAYPM53 (Delay Getting PMED For COVID R3/1, R4/2, R5/3) indicate with a value of ‘1’ (Yes) that the person was delayed in receiving that care during the pandemic; a value of ‘2’ (No) for these variables indicates that the person was not delayed in receiving that type of care during the pandemic.

COVID-19 Vaccination Status

The CV section also gathers information regarding vaccination and booster shots ever received for COVID-19 for all members of the RU. CVVACCINE42 and CVVACCINE53 represent round-specific measures of ever having received the COVID-19 vaccination. Sample members who were reportedly ever vaccinated as of Round 4/2 (CVVACCINE42=1) had CVVACCINE53 coded ‘1’ (Ever vaccinated) even though they were not asked in Round 5/3 whether they were ever vaccinated. BOOSTERSHOT53 was collected only for Panel 23 Round 9, Panel 24 Round 7, Panel 25 Round 5, and Panel 26 Round 3, and it indicates whether the person had ever received a COVID-19 vaccine booster shot before the end of the reference period.

2.5.9 Employment Variables (EMPST31-RTPLN53H)

Employment questions were asked of all persons 16 years and older at the time of the interview. Employment variables consist of person-level indicators such as employment status and job-

related variables such as hourly wage. All job-specific variables refer to a person’s current main job. The current main job, defined by the respondent, indicates the main source of employment.

Historically, most employment variables pertain to the interview date for Rounds 1-4, and to December 31 of the delivery year for Round 5 of a second year panel. In 2021, Employment variables have been constructed to reflect responses from additional panels. Panel 24 was fielded for a third year and includes responses from Round 5, Round 6, and Round 7 interviews. Panel 23 was fielded for a fourth year and includes responses from Round 7, Round 8, and Round 9 interviews.

In 2021, Panel 25 Round 3, Panel 24 Round 5, and Panel 23 Round 7 were fielded as cross-year rounds where respondents were asked to provide information about the reference period between the prior interview date in 2020 (Rounds 2, 4, and 6 respectively) and the current round interview date in 2021. Panel 26 Round 3 and Panel 24 Round 7 were also fielded as cross-year rounds where respondents were asked to provide information about the reference period between the prior interview date in 2021 (Rounds 2 and 6 respectively) and current round interview date (occurring in 2022).

In contrast, Panel 23 Round 9 and Panel 25 Round 5 were fielded as 2021 terminal rounds where respondents were asked to provide relevant information between the prior interview date in 2021 (Rounds 8 and 4, respectively) and 12/31/2021.

The following table summarizes reference periods used in construction of Employment variables in this delivery:

Panel/Round	Reference Period (construction)	EM Variable Ref Year(s)
Panel 23		
<i>Round 7</i>	Round 6 2020 intvw thru Round 7 2021 intvw date	2020 - 2021
<i>Round 8</i>	Round 7 2021 intvw thru Round 8 2021 intvw date	2021
<i>Round 9</i>	Round 8 2021 intvw thru Dec 31, 2021	2021
Panel 24		
<i>Round 5</i>	Round 4 2020 intvw thru Round 5 2021 intvw date	2020 - 2021
<i>Round 6</i>	Round 5 2021 intvw thru Round 6 2021 intvw date	2021
<i>Round 7</i>	Round 6 2021 intvw thru Round 7 2022 intvw date	2021 - 2022
Panel 25		
<i>Round 3</i>	Round 2 2020 thru Round 3 2021 intvw date	2020-2021
<i>Round 4</i>	Round 3 2021 intvw thru Round 4 2021 intvw date	2021
<i>Round 5</i>	Round 4 2021 intvw thru Dec 31, 2021	2021
Panel 26		
<i>Round 1</i>	Jan 1, 2021 thru Round 1 2021 intvw date	2021
<i>Round 2</i>	Round 1 2021 intvw thru Round 2 2021 intvw date	2021
<i>Round 3</i>	Round 2 2021 intvw thru Round 3 2022 intvw date	2021-2022

No additional adjustments were necessary for 2021 variables, however, users combining 2021 data with 2020 MEPS data should refer to 2020 documentation to fully understand adjustments that were made for Panel 23 variables in Round 5 and Round 6 in the 2020 file.

Variable naming protocol is consistent with all prior years. Historically, round dates have been indicated by two numbers following the variable name; the first number representing the round for second panel persons (Panel 25), the second number representing the round for first panel persons (Panel 26). For example, EMPST31 refers to employment status on the Round 3 interview date for Panel 25 persons and employment status on the Round 1 interview date for Panel 26 persons. In 2021, a third year panel and a fourth year panel are included in each of the '31', '42', and '53' variables, but round numbers of these panels (5/6/7 and 7/8/9 respectively) will not be included in the variable name. For example, the 2021 version of EMPST31 (noted above) will also include employment status on the Round 7 interview date for Panel 23 persons and the employment status on the Round 5 interview date for Panel 24 persons. All employment variables reflect the inclusion of the third year of Panel 24 and the fourth year of Panel 23. Panel 23 Round 7 and Panel 24 Round 5 information is contained in the '31' variables, Panel 23 Round 8 and Panel 24 Round 6 information is contained in the '42' variables, and Panel 23 Round 9 and Panel 24 Round 7 information is contained in the '53' variables. (Some users might find it helpful to think of these variables as (75)31, (86)42, and (97)53, even though the naming convention remains as 31, 42, and 53).

With the exception of some health insurance and wage variables, no attempt has been made to logically edit any employment variables. When missing, values were imputed for certain persons' hourly wages. Due to confidentiality concerns, hourly wages greater than or equal to \$105.77 were top-coded to -10 and the number of employees variable was top-coded at 500. With the exception of a variable indicating whether the employer has more than one location (MORE31, MORE42, MORE53), all employer-specific variables on the Full Year Public Use Release file refer to the specific establishment that is the location of a person's current main job.

For users interested in additional jobs (i.e. current miscellaneous, former main job, and others) or additional details about the current main job, please refer to the Jobs Public Use release file for the current delivery year.

The MEPS employment section used dependent interviewing in Rounds 2 through 9. If employment status and certain job characteristics did not change from the previous round, as identified in the Review of Employment (RJ) section, the respondent was skipped through the main employment section. A code of "Determined in Previous Round" (-2) is used to indicate that the information in question was obtained in a previous round. For example, if HRWG42X (Round 8 interview date hourly wage for Panel 23 persons, or Round 6 interview date hourly wage for Panel 24 persons, or Round 4 interview date hourly wage for Panel 25 persons, or Round 2 interview date hourly wage for Panel 26 persons) is coded as "Determined in Previous Round" (-2), it means that hourly wage was collected in a previous round. In this case, users would need to refer to HRWG31X (Round 7 interview date hourly wage for Panel 23 persons, or Round 5 interview date hourly wage for Panel 24 persons, Round 3 interview date hourly wage for Panel 25 persons, or Round 1 interview date hourly wage for Panel 26 persons) to obtain the value for HRWG42X. The "-2" value for HRWG42X indicates that the person was skipped past the hourly wage question at the time of the Round 8/6/4/2 interview. The same coding applies to

HRWG53X when a person was skipped past the Round 9/7/5/3 interview. Note that users may find a positive value in the HRWG31X (Round 7/5/3/1 hourly wage) or they may find an “Inapplicable” code of -1. Unlike HRWG42X and HRWG53X, HRWG31X does not contain -2 values. For persons skipped in Round 7/5/3/1, the prior year hourly wage value is used to populate HRWG31X. The prior year round from which the wage was collected for such cases can be found in the RNDGLG31 variable. These cases will be discussed in more detail below.

To determine who should be skipped through various employment questions, certain information, such as employment status, had to be asked in every round. Note that “-2” codes do not apply to questions asked in every round, like employment status. Additionally, information on whether the person currently worked at more than one job (MORJOB) or whether the person held health insurance from a current main employer (HELDX) was asked in every round, and, therefore, those variables also have no “-2” codes.

For (a) Panel 23 persons who have a current main job in Round 7 that continued from a job first reported in Round 1 or 2 of 2018, (b) Panel 23 persons who have a current main job in Round 7 that continued from a job first reported in Round 3 or 4 of 2019, (c) Panel 23 persons who have a current main job in Round 7 that continued from a job first reported in Round 5 or Round 6 of 2020, (d) Panel 24 persons who have a current main job in Round 5 that continued from a job first reported in Round 1 or 2 of 2019, (e) Panel 24 persons who have a current main job in Round 6 that continued from a job first reported in Round 3 or Round 4 of 2020, or (f) Panel 25 persons who have a current main job in Round 3 that continued from a job first reported in Round 1 or Round 2 of 2020, the “-2” code is not used. This is because prior year employment variables do not reside on the current 2021 Full Year Release file, and are therefore not easily accessible for users (and in some cases, the data could be impossible to obtain). Instead, the values for the variables resulting from skipped questions are copied from the appropriate prior year file (2018, 2019, or 2020) to the 2021 Full Year Public Use Release ‘31’ variable, depending on the round in which the job first became the current main job:

- the Panel 23 Round 1 or 2 constructed variable from the 2018 Full Year Release file, or
- the Panel 23 Round 3 or 4 constructed variable from the 2019 Full Year Release file, or
- the Panel 23 Round 5 or 6 constructed variable from the 2020 Full Year Release file or
- the Panel 24 Round 1 or 2 constructed variable from the 2019 Full Year Release file.
- the Panel 24 Round 3 or 4 constructed variable from the 2020 Full Year Release file or
- the Panel 25 Round 1 or 2 constructed variable from the 2020 Full Year Release file.

The accompanying 2021 variable RNDFLG31 indicates the round from which these data were collected. For example, if the Panel 25 person has a Round 3 current main job that continues

from Round 2 and was first reported as the current main job in Round 2, HRWG31X in the 2021 Full Year Public Use release will be a copy of the HRWG42X variable from the 2020 Full Year Public Use release, and RNDFLG31 in the 2021 Full Year Public Use release will be “2”, indicating the round in which the job was first reported as the current main job.

Employment Status (EMPST31, EMPST42, and EMPST53)

Employment status was asked for all persons age 16 or older. Allowable responses to the employment status questions were as follows:

- “currently employed” if the person had a job at the interview date;
- “has a job to return to” if the person did not work during the reference period but had a job to return to as of the interview date;
- “employed during the reference period” if the person had no job at the interview date but did work during the round;
- “not employed with no job to return to” if the person did not have a job at the interview date, did not work during the reference period, and did not have a job to which they could return.

These responses were mutually exclusive. A current main job was defined for persons who either reported that they were currently employed and identified a current main job or who reported and identified a job to return to. Therefore, job-specific information such as hourly wage exists for persons not presently working at the interview date but who have a job to return to as of the interview date.

The analyst should note that there are cases where $EMPST_{rr} = 1$ or 2 (has current job or job to return to) where DDNWRK21 contains a positive value, indicating the number of times the person lost a half-day or more from work. This is because the responses to the Disability Days questions in the survey are independent of the responses to the employment questions.

Data Collection Round for Round 7, 5, 3, or 1 CMJ (RNDFLG31)

As mentioned above, for (a) a Panel 25 person with a Round 3 current main job (CMJ) that is a continuation CMJ from Round 1 or Round 2, (b) a Panel 24 person with a Round 5 CMJ that is a continuation CMJ from Round 1, 2, 3, or 4, or (c) a Panel 23 person with a Round 7 CMJ that is a continuation CMJ from Round 1, 2, 3, 4, 5, or 6, the value for most ‘31’ variables will be copied forward from the 2018, 2019, or 2020 Full Year Public Use release from the variable representing the round in which the job was first reported as the CMJ. Therefore, for persons in Panel 23, Panel 24, or Panel 25, RNDFLG31 indicates the 2018, 2019, or 2020 round in which the Round 7, Round 5, or Round 3 CMJ was first reported as the CMJ and provides a timeframe for the reported wage information and other job details. RNDFLG31 is used with many ‘31’ variables to indicate the round on which the reported information is based.

RNDFLG31 is set to “Inapplicable” (-1) for persons in any panel who are under age 16 or who do not have a CMJ in Panel 23 Round 7, Panel 24 Round 5, Panel 25 Round 3, or Panel 26 Round 1. For persons who are part of Panel 23, RNDFLG31 is also set to “Inapplicable” (-1) if the person is out-of-scope in the 2021 portion of Round 7. For persons who are part of Panel 24, RNDFLG31 is also set to “Inapplicable” (-1) if the person is out-of-scope in the 2021 portion of Round 5. For persons who are part of Panel 25, RNDFLG31 is also set to “Inapplicable” (-1) if the person is out-of-scope in the 2021 portion of Round 3. For persons who are part of Panel 26, RNDFLG31 is also set to “Inapplicable” (-1) if the person is out-of-scope in Round 1. Values for RNDFLG31 are set as follows:

- 1 continuing Panel 23 Round 7/Panel 24 Round 5/Panel 25 Round 3 CMJ reported first in Round 1, or Panel 26 Round 1 CMJ newly reported as current main in Round 1
- 2 continuing Panel 23 Round 7/Panel 24 Round 5/Panel 25 Round 3 CMJ reported first in Round 2
- 3 continuing Panel 23 Round 7/Panel 24 Round 5 CMJ reported first in Round 3 or Panel 25 Round 3 CMJ newly reported as current main in Round 3
- 4 continuing Panel 23 Round 7/Panel 24 Round 5 CMJ reported first in Round 4
- 5 continuing Panel 23 Round 7 CMJ reported first in Round 5 or Panel 24 Round 5 CMJ newly reported as current main in Round 5
- 6 continuing Panel 23 Round 7 CMJ reported first in Round 6
- 7 Panel 23 Round 7 CMJ newly reported as current main in Round 7
- 15 Panel 23 Round 7/Panel 24 Round 5 CMJ/Panel 25 Round 3 CMJ is a continuation CMJ (wage information and other details were not collected in Round 7/Round 5/Round 3) but the Round 6/Round 4/Round 2 CMJ record either does not exist or is not the same job. This setting applies even in cases where there is a corresponding Round 1, 2, 3, 4, or 5 CMJ for Panel 23, Round 1, 2, or 3 CMJ for Panel 24 or Round 1 CMJ for Panel 25. This can occur in rare instances because corrections made to a person’s record in a current file cannot be made to that record in an earlier file due to database processing constraints. Corrections are made based on respondent comments in subsequent rounds that affect employment information previously reported. Users may refer to previously released Jobs Public Use release files to review rosters as follows:
 - 2018 Jobs Public Use release file - for Panel 23 Round 1 through 3 rosters or
 - 2019 Jobs Public Use release file - for Panel 23 Round 3 through 5, Panel 24 Round 1 through 3 rosters or

- 2020 Jobs Public Use release file - for Panel 23 Round 5 through 7, Panel 24 Round 3 through 5, or Panel 25 Round 1 through 3 rosters

Self-Employed (SELFCM31, SELFCM42, and SELFCM53)

Information on whether an individual was self-employed at the current main job was obtained for all persons who reported a current main job. If an individual reports that they are self-employed at their current main job, they are also asked to identify whether the self-employed business was incorporated, a proprietorship, or a partnership (BSNTY31, BSNTY42, BSNTY53). These questions are not asked of individuals who were not self-employed and, as a result, individuals who are not self-employed are coded with “Inapplicable” (-1).

Alternatively, there are several variables that are only constructed for wage earners (not self-employed). These include benefits, employment characteristics, and hourly wage variables (covered in the following two sections). As noted below, self-employed individuals are coded with “Inapplicable” (-1) for benefits, employment characteristics, and hourly wage variables.

Benefits and Employment Characteristics (PAYDR31/42/53, SICPAY31/42/53, PAYVAC31/42/53, RETPLN31/42/53, MORE31/42/53, JOBORG31/42/53)

Several variables are constructed only for individuals who report not being self-employed at their current main job. These individuals are asked questions to indicate whether the establishment reported as the main source of employment offered any of the following benefits:

- Paid leave to visit a doctor (PAYDR31, PAYDR42, and PAYDR53)
- Paid sick leave (SICPAY31, SICPAY42, and SICPAY53)
- Paid vacation (PAYVAC31, PAYVAC42, and PAYVAC53)
- Pension plan (RETPLN31, RETPLN42, and RETPLN53)

They are also asked information on whether the firm had more than one business location (MORE31, MORE42, MORE53) and whether the establishment was a private for-profit, nonprofit, or a government entity (JOBORG31, JOBORG42, JOBORG53). For persons who were self-employed at their current main job, all of the variables detailed in this section were coded as “Inapplicable” (-1).

Hourly Wage (HRWG31X, HRWG42X, HRWG53X), Wage Update Variable (DIFFWG31, DIFFWG42, DIFFWG53), and Updated Hourly Wage (NHRWG31, NHRWG42, NHRWG53)

Hourly wage was constructed for all persons who reported a current main job that was not self-employment (SELFCM). HRWG31X, HRWG42X, and HRWG53X provide the wage amount reported initially for a person’s current main job. If a current main job continues into subsequent

rounds DIFFWG31, DIFFWG42, and DIFFWG53 indicate if the wage has changed since the previous round. If the job continues and there is a different wage at that job, NHRWG31, NHRWG42, and NHRWG53 indicate the new wage amount. The initial hourly wage variables (HRWG31X, HRWG42X, HRWG53X) on this file should be considered along with their accompanying variables - HRHOW31, HRHOW42, and HRHOW53 - which indicate how the respective round hourly wage was constructed.

Hourly wage could be derived, as applicable, from a large number of source variables. In the simplest case, hourly wage was reported directly by the respondent. For other persons, construction of the hourly wage was based upon salary, the time period on which the salary was based, and the number of hours worked per time period. If the number of hours worked per time period was not available, a value of 40 hours per week was assumed, as identified in the HRHOW variable.

To assist interviewers during collection of wage amounts, CAPI prompts the respondent to confirm wages reported in the Employment Wage section if a wage amount falls outside a specified wage range. Ranges vary depending on the unit of pay as follows:

Unit of Pay	Wage Range
Per year	\$5,000.00 - \$200,000.00
Per month	\$375.00 - \$20,000.00
Per 2-week period	\$150.00 - \$10,000.00
Per week	\$75.00 - \$5,000.00
Per day	\$10.00 - \$750.00
Per hour	\$1.00 - \$125.00

Where there was insufficient information available for calculating the initial hourly wage, the initial hourly wage variables HRWG31X, HRWG42X, and HRWG53X were imputed using a weighted sequential hot-deck procedure for individuals who reported a current main job (and were not self-employed) but did not know their wage or refused to report a wage. Hourly wage for persons for whom employment status was not known was coded as “Cannot be Computed” (-15). Additionally, wages were imputed for wage earners who reported a wage range instead of a specific wage value. For each of these persons, a value was imputed from other persons on the file who did report a specific value that fell within the reported range. Wages from 2018, 2019, 2020, and 2021 were eligible donors in this process, expanding the donor pool to cover four years instead of the typical two years. The expansion of the donor pool to use four years of donors instead of two allowed AHRQ to maintain a similar sized donor pool to prior releases - but it does mean that some recipients are assigned a donor wage from four years prior.

The variables HRWGIM31, HRWGIM42, and HRWGIM53 identify persons whose wages were imputed. Note that wages were imputed only for persons with a positive person and/or positive family weight.

The variables DIFFWG31, DIFFWG42, and DIFFWG53 indicate whether a person’s wage amount was different in the current round (from the previous round) at a continuing, current main job. NHRWG31, NHRWG42, and NHRWG53 contain the updated wage amount in cases

where a person indicates a change in wages (DIFFWG = 1). While the question regarding wage changes pertains to the primary wage at the main job, occasionally respondents update a person's supplemental wage at the main job. In these cases, users should note that the HRWG31X, HRWG42X, HRWG53X variable may contain the same value as the NHRWG31, NHRWG42, NHRWG53 variable.

For all Panel 26 Round 1 persons, DIFFWG31 and NHRWG31 are set to "Inapplicable" (-1) because this was the first round that wages could be reported for those persons. In Rounds 2 through 9 for all panels, no imputation was performed on NHRWG31, NHRWG42, and NHRWG53. Instead, where an updated wage amount is "Don't Know" (-8) or is "Refused" (-7), NHRWG31, NHRWG42, and/or NHRWG53 is set to "Cannot be Computed" (-15). For persons whose hourly wage variable HRWG31X, HRWG42X, and/or HRWG53X was imputed and the respondent provides an updated wage amount in a subsequent round, the new wage, NHRWG31, NHRWG42, and/or NHRWG53, is not presented. Instead, NHRWG31, NHRWG42, and/or NHRWG53 is set to "Initial Wage Imputed" (-13) to indicate that the initial HRWG31X, HRWG42X, and/or HRWG53X was imputed. Users are able to access the value reported for updated wage for these jobs by referring to the 2021 Jobs Public Use release file.

In 2021, wage information is logically edited for consistency using established rules and guidance from AHRQ. Outliers are checked for persons who report a wage change and the new reported wage (a) is substantially different from prior wage (change $\geq 100\%$), (b) is no different than prior wage, (c) is low in value ($\$0 < \text{wage} < \1) or, (d) has a value higher than prior year's top code value. There are numerous sources for these types of errors, including keystroke or respondent error. In 2021, approximately 100 wages were reviewed per panel, resulting in approximately 90 HRWGrrX/NHRWGrr wage variable edits (overall).

Users should note that outlier edits were not performed in 2020 and therefore should be mindful when using the wage variables, especially when comparing 2020 wages to wages in other data years. To help users identify cases that would have been reviewed (but not necessarily edited) in this process, the 2020 data includes wage outlier flag variables, OUTFLAGrr. These round-specific wage outlier flag variables OUTFLAG31, OUTFLAG42, and OUTFLAG53 indicated that a person's updated wage at the current main job would have been programmatically selected for review during the 2020 wage outlier editing process (but not necessarily edited). Although the OUTFLAGrr variables only appear on the 2020 file, they could be relevant to continuing wages on the 2021 file that were first reported in 2020. More information on these variables may be found in MEPS HC-224: 2020 Full Year Consolidated Data file documentation. OUTFLAGrr variables were not constructed for the 2021 Population Characteristics file since outlier reviews were performed.

For reasons of confidentiality, the hourly wage variables were top-coded. A value of -10 indicates that the hourly wage was greater than or equal to \$105.77. The wage top-code process uses the highest calculated wage for an individual regardless of whether it was reported in the HRWG31X, HRWG42X, and HRWG53X variable or the NHRWG31, NHRWG42, and NHRWG53 variable. All wages for a person were top-coded if any wage variable was at or above the top-code amount.

In order to protect the confidentiality of persons across deliveries, the same top-code amount used in this 2021 Full-Year Consolidated file was also applied to the 2021 Jobs file. Because a person can have other jobs besides a current main job which are included in the corresponding 2021 Jobs file, wages at these other jobs were reviewed in the top-coding process. In some cases for these persons, wages reported at the current main job were below the top-code amount while the wage at another job had to be top-coded. To further protect the confidentiality of such persons across deliveries, wages reported at all jobs in the 2021 Jobs file were top-coded and the wages at their current main job (HRWG31X, HRWG42X, HRWG53X, NHRWG31, NHRWG42, and NHRWG53) included in this 2021 Full-Year Consolidated file were also top-coded.

Health Insurance (HELD31X, HELD42X, HELD53X, OFFER31X, OFFER42X, OFFER53X, CHOIC31, CHOIC42, CHOIC53, DISVW31X, DISVW42X, DISVW53X, OFREMP31, OFREMP42, OFREMP53)

There are several employment-related health insurance measures included in this release: health insurance held at a current main job (HELD31X, HELD42X, HELD53X), health insurance offered through a current main job (OFFER31X, OFFER42X, OFFER53X), health insurance offered to anyone through the current main job employer (OFREMP31, OFREMP42, OFREMP53), and choice of health plans available through the current main job (CHOIC31, CHOIC42, CHOIC53). This collection of variables reflects the insurance status in the current round. The variables are logically edited for consistency for each round.

MEPS asks if the person holds health insurance through the current main job (HELDX) in the first round in which a person is reported as having that job. If the person does not hold health insurance at the job, then a follow-up question is asked as to whether the person was offered insurance but declined coverage (OFFERX). If the person neither holds nor was offered health insurance at the job, then an additional question is asked to determine whether any other employees at the current main job were offered health insurance (OFREMP). If the person either holds insurance from the job or was offered insurance at the job, then an additional question is asked to determine whether a choice of health plans is available at the job (CHOIC). Prior to Panel 23 Round 9, Panel 24 Round 7, Panel 25 Round 5, and Panel 26 Round 3, in cases where HELDX or OFFERX were “Refused” (-7) or “Don’t Know” (-8), CHOIC was also coded -7/-8, even though the question that populates CHOIC was not asked. As of Panel 23 Round 9, Panel 24 Round 7, Panel 25 Round 5, and Panel 26 Round 3, CHOIC is now coded “Cannot be Computed” (-15) for cases where HELDX or OFFERX were “Refused” (-7) or “Don’t Know” (-8) to reflect that a value cannot be calculated for CHOIC as a result of skip patterns.

In the rounds after the job is first reported, the Review of Jobs (RJ) section has the same series of questions with one exception, it does not ask whether there is a choice of health insurance plans at an employer. Choice of health insurance plan is only asked in the round the job was first reported (in the Employment section).

In rounds after the job is first reported, a “held” question (whether a person now holds health insurance through the employer) is asked in the Review of Jobs section. This is to determine if there has been any change in coverage.

RJ70 (HELDX) is asked if insurance was offered, but not taken by the employee, when the job was first reported and no coverage has been reported since the initial round.

RJ80 (HELDX) is asked where:

- insurance coverage through the job ended in a prior round or
- insurance coverage through the job was never reported and the person was not offered insurance through the job in the round the job was first reported or
- the respondent disavows insurance coverage in the Health Insurance section despite having previously indicated insurance coverage through the job in the Employment section of the interview or
- beginning in Panel 23 Round 9, Panel 24 Round 7, Panel 25 Round 5, and Panel 26 Round 3, persons who reported new employer-sponsored health insurance coverage in the prior round through the CMJ but the insurance covered the person for only part of that round (see detailed explanation below).

MEPS then includes several clarifying questions regarding health insurance status and availability to the jobholder through an employer. Where the person does not report, does not know, or refuses to indicate the insurance status at RJ70, or reports no coverage at RJ80, the respondent is asked if the person was offered insurance (OFFERX).

Lastly, when a respondent indicates that the jobholder of a reviewed job neither holds nor was offered health insurance at the job, the respondent is asked if *any other* employees at the job were offered health insurance (OFREMP).

If a person does hold insurance through their job, then that person is not asked the offer question and the OFFERX variable is automatically set to “Yes” (1). Data users should note that OFREMP is automatically set to 1 in cases where the jobholder has health insurance coverage through the job (HELDX=1) or in cases where health insurance is offered to the employee at their job (OFFERX=1).

Responses in the Employment and Review of Jobs sections for health insurance held were recoded to be consistent with the variables in the Health Insurance section of the survey.

For persons who responded in the Employment section or Review of Jobs section that they held health insurance coverage through the employer but then disavowed the coverage in the Health Insurance section, MEPS includes follow-up questions regarding whether health insurance is offered (either to the employee or any other employee depending on responses to questions) and whether more than one plan is available. This information is used in an editing process whereby responses to these variables in the Health Insurance section are transferred into the Employment section or Review of Jobs section. Consequently, more information is now available on the MEPS file in the OFFERX, OFREMP, and CHOIC variables. Consistent with prior years, the round-specific flag variable DISVWX continues to be constructed and reflects the disavowal at the current main job in the round.

Beginning Panel 23 Round 9, Panel 24 Round 7, Panel 25 Round 5, and Panel 26 Round 3, two CAPI changes impact how insurance information is collected in the Employment and Review of Jobs sections.

1. RJ80 is asked if a person reported new health insurance in the prior round but that coverage was not active at the interview date, that is, a response of “No” (2) at HQ01 “Was {PERSON} covered the whole time from {START DATE} until {END DATE}” and at HQ02 “Is {PERSON} covered now?”. This CAPI change means that more persons could be asked if the person was offered insurance or if other employees were offered insurance at the employer establishment. The result is that HELDX, OFFERX, OFREMP could have fewer -15 values.
2. If a job holder has insurance at the employer (“Yes” (1) at EM660) and that person belongs to a union (“Yes” (1) at EM700), respondents are asked to indicate if the health insurance is from the employer/business or the union at EM710. Either or both establishments may be the source of insurance. Prior to the CAPI change, where both establishments were selected, two sets of private insurance coverage were created in the Health Insurance section. Now, only the primary source of private insurance coverage will be created in the Health Insurance section.

Now that the response options have changed, respondents are required to identify the *primary* source - either employer/business or the union - if the respondent indicates both provide insurance.

1 EMPLOYER

2 UNION

3 BOTH EMPLOYER AND UNION (EMPLOYER IS PRIMARY)

4 BOTH EMPLOYER AND UNION (UNION IS PRIMARY)

The result is that persons who report insurance via both union and employer sources in the Employment section will no longer have the secondary source insurance coverage recorded in the Health Insurance section. However, respondents continue to have the opportunity to report any additional private coverage in the Health Insurance section at HX190/HX200.

Hours (HOUR31, HOUR42, HOUR53)

The hours variables refer to usual hours worked per week at the current main job. Note that, in cases where the respondent estimated hours worked per week at 35 hours or more, HOUR31, HOUR42, and HOUR53 were set to ‘40.’

Temporary (TEMPJB31, TEMPJB42, TEMPJB53) and Seasonal (SSNLJB31, SSNLJB42, SSNLJB53) Jobs

The temporary job variables (TEMPJB31, TEMPJB42, TEMPJB53) indicate whether a *newly reported* current main job lasts for only a limited amount of time or until the completion of a project.

The seasonal job variables (SSNLJB31, SSNLJB42, SSNLJB53) indicate whether the *newly reported* current main job is only available during certain times of the year. SSNLJB is “Yes” (1) if the job is only available during certain times of the year; SSNLJB is “No” (2) if the job is year round. Teachers and other school personnel who work only during the school year are considered to work year round.

Both variables are set on current main jobs whether a person is self-employed or not. These questions are asked only in the round the job is newly reported. Consequently, in rounds following the initial report, a code of “Determined In Previous Round” (-2), is used to indicate that the information in the question was obtained in a previous round. This differs from some previous files where both questions were asked in each round and -2 was not an allowed value. Analysts using either of these variables over multiple years of MEPS should refer to documentation for each year to assure consistency for the variable.

Number of Employees (NUMEMP31, NUMEMP42, NUMEMP53)

NUMEMP indicates the number of employees at the location of the person’s current main job. Due to confidentiality concerns, this variable has been top-coded at 500 or more employees. For respondents who do not know the specific number of employees at the establishment, a categorical question is offered as an alternative. In these cases, a numerical value for NUMEMP is constructed using a median estimated size calculated from donors within the reported categorical range. As always, median values may vary across panels/rounds since calculations are panel/round specific. One noticeable difference in 2021, however, was on medians calculated for NUMEMP31 in the highest estimated range (101-500 employees). Where Panel 26 had a median establishment size of 220, Panel 25 had a median establishment size of 200, Panel 24 had a median establishment size of 300, and Panel 23 had a median establishment size of 150. Otherwise, differences were generally less pronounced.

CAPI does not accept an establishment size value of 0 to indicate the total number of employees working at a **self-employed business**. Where a person is **not self-employed** at a job, an establishment size of 0 is allowed. NUMEMP is set to “Cannot be Computed” (-15) for these cases.

Other Employment Variables

Information about industry and occupation types for a person’s current main job at the interview date is also contained in this release. Based on verbatim text fields collected during the interview, numeric industry and occupation codes are assigned by trained coders at the Census Bureau. The coders used 2007 Census Industry and 2010 Census Occupation Coding schemes,

which were developed for the Bureau's Current Population Survey and American Community Survey. Users should note that coding schemes are comparable for the FY 2010 through FY 2021 data files. Earlier versions of Census coding schemes were used on files prior to FY 2010.

Current main jobs were initially coded at the 4-digit level for both industry and occupation. Then, for confidentiality reasons, these codes were condensed into broader groups for release on the file. INDCAT31, INDCAT42, and INDCAT53 represent the condensed industry codes for a person's current main job at the interview date. OCCCAT31, OCCCAT42, and OCCCAT53 represent the condensed occupation codes for a person's current main job at the interview date.

This release incorporates crosswalks showing how the detailed 2007 Census industry codes (Appendix 1) and 2010 Census occupation codes (Appendix 2) were collapsed into the condensed codes on the file. The schemes used in this file can be linked directly to the 2007 North American Industry Code System (NAICS) and the 2010 Standard Occupation Code scheme (SOC) by going to the [U.S. Census Bureau website](#) where a variety of additional crosswalks is also available.

Information indicating whether a person belonged to a labor union (UNION31, UNION42, and UNION53) is also contained in this release.

The month and year that the current main job started for Rounds 7, 8, and 9 of Panel 23; Rounds 5, 6, and 7 of Panel 24; Rounds 3, 4, and 5 of Panel 25; and Rounds 1, 2, and 3 of Panel 26 are provided in this release (STJBMM31, STJBYY31, STJBMM42, STJBYY42, STJBMM53, and STJBYY53). In Full Year 2021, STJBYY31, STJBYY42, and STJBYY53 are bottom coded to a value of '1951' to preserve age confidentiality. This value is calculated by taking the delivery year of 2021 and subtracting the age top code value of 85, then adding back 15, the age of a person in the year before entering the work force as defined in MEPS. Thus, the bottom code value will be different in each delivery year. Because a current main job that continues into Panel 23 Round 7, Panel 24 Round 5, and Panel 25 Round 3 may have been reported in a previous delivery year, bottom code values vary for each panel. Therefore, the possible STJBYY31 values are as follows:

- For Panel 23 Round 7 the values are 1948 - 2021 since the job may have first been reported in 2018
- For Panel 24 Round 5 the values are 1949 - 2021 since the job may have first been reported in 2019
- For Panel 25 Round 3 the values are 1950 - 2021 since the job may have first been reported in 2020
- For Panel 26 Round 1 the values are 1951 - 2021.

There are two measures included in this release that relate to a person's work history over a lifetime. One indicates whether a person ever retired from a job as of the Round 9 interview date for Panel 23 persons, or as of the Round 7 interview date for Panel 24 persons, or as of the Round 5 interview date for Panel 25 persons, or as of the Round 3 interview date for Panel 26

persons (EVRETIRE). The other indicates whether a person ever worked for pay as of the Round 9 interview date for Panel 23 persons, as of the Round 7 interview date for Panel 24 persons, as of the Round 5 interview date for Panel 25 persons, or as of the Round 3 interview date for Panel 26 persons (EVRWRK). The latter was asked of everyone who indicated that they were not working as of the round interview date. Therefore, anyone who indicated current employment or who had a job during any of the previous or current rounds was skipped past the question identifying whether the person ever worked for pay. These individuals were coded as “Inapplicable” (-1). All persons who ever reported a job and were 55 years or older as of the round interview date were asked if they “ever retired”. Since both of these variables are not round specific, there are no “Determined in Previous Round” (-2) codes.

This release contains variables indicating the main reason a person did not work since the start of the reference period (NWK31, NWK42, and NWK53). If a person was not employed at all during the reference period (at the interview date or at any time during the reference period) but was employed some time prior to the reference period, the person was asked to choose from a list the main reason they did not work during the reference period.

The “Inapplicable” (-1) category for the NWK variables includes:

- Persons who were employed during the reference period;
- Persons who were not employed during the reference period and who were never employed;
- Persons who were out-of-scope the entire reference period and;
- Persons who were less than 16 years old.

A measure of whether an individual had more than one job on the round interview date (MORJOB31, MORJOB42, and MORJOB53) is provided on this release. In addition to those under 16 and those individuals who were out-of-scope, the “Inapplicable” (-1) category includes those who did not report having a current main job. Because this is not a job-specific variable, there are no “Determined in Previous Rounds” (-2) codes.

This release contains a variable indicating if a current main job changed between the seventh and eighth rounds for Panel 23 persons, between the fifth and sixth rounds for Panel 24 persons, between the third and fourth rounds for Panel 25 persons, or between the first and second rounds for Panel 26 persons (CHGJ3142). It also contains a variable indicating if a current main job changed between the eighth and ninth rounds for Panel 23 persons, between the sixth and seventh rounds for Panel 24 persons, between the fourth and fifth rounds for Panel 25 persons, or between the second and third rounds for Panel 26 persons (CHGJ4253). In addition to the “Inapplicable” (-1), “Refused” (-7), “Don’t Know” (-8), and “Cannot be Computed” (-15) categories, the change job variables were coded to represent the following:

- 1 person left previous round current main job and now has a new current main job;
- 2 person still working at the previous round’s current main job but, as of the new round, no longer considers this job to be the current main job and defines a new

current main job (previous round’s current main job is now a current miscellaneous job);

- 3 person left previous round’s current main job and does not have a new job;
- 4 person did not change current main job.

Finally, this release contains the reason given by the respondent for the job change (YCHJ3142 and YCHJ4253). The reasons for a job change were listed in the CAPI questionnaire and a respondent was asked to choose the main reason from this list. Beginning Panel 23 Round 9, Panel 24 Round 7, Panel 25 Round 5, and Panel 26 Round 3, in addition to those out-of-scope, those under 16, those not having a current main job, and workers who did not change jobs, the “Inapplicable” (-1) category for YCHJ3142 and YCHJ4253 now also includes workers who continue to work at the main job but no longer consider it their main job (CHGJrrrr = Changed CMJ/Previous CMJ is Now Current Miscellaneous job” (2)). These persons did not leave the job and therefore were not asked why they left a job. Prior to this change, these persons were set to “Cannot be Computed” (-15).

Employment Variables Imputed for Missing Values (EMPST31H - RTPLN53H)

To further assist analysts, a series of fully-imputed employment variables is available on the Full Year Consolidated Data File (FY PUF). For the years 2000 to 2013 these variables can be found on MEPS HC-131 (MEPS Employment Variables 2000-2013). For the years 2014 and beyond, these variables can be found in the FY PUF for each data year. The fully-imputed variables in this file are developed from the constructed Employment variables in the 2021 FY PUF.

Observations for these employment variables with values of “Refused” (-7), “Don’t Know” (-8), or “Cannot be Computed” (-15) were imputed using weighted sequential hot-decking. The imputations were performed separately for each MEPS HC panel across the nine survey rounds of the MEPS. First, all missing values of a given variable were imputed for Round 1. If a person remained in the same job in Round 2, and the MEPS questionnaire did not ask for updated job information (i.e., if the variable on the FY PUF was coded as -2), then the value for that variable was pulled forward from Round 1 to Round 2 (including values that had been imputed in Round 1). After pulling values forward from Round 1, any remaining missing values were imputed for Round 2. This process was repeated for Rounds 3-9. For Panel 26, the imputed employment variables use constructed employment variables from Round 1, 2 and 3 data on this file. However, in order to impute employment variables for Panel 23, Panel 24, and Panel 25, data from the current and prior delivery years are required as follows:

	MEPS HC-209	MEPS HC-216	MEPS HC-224	MEPS HC-233
	<u>2018</u>	<u>2019</u>	<u>2020</u>	<u>2021</u>
Panel 23	Round 1 - Round 2	Round 3 - Round 4	Round 5 - Round 6	Round 7 - Round 9
Panel 24		Round 1 - Round 2	Round 3 - Round 4	Round 5 - Round 6
Panel 25			Round 1 - Round 2	Round 3 - Round 5
Panel 26				Round 1 - Round 3

For Panel 23 Round 7, Panel 24 Round 5, and Panel 25 Round 3 that cross the calendar year from 2020 to 2021, the most up-to-date version of data in the round are used. While MEPS HC-224 also has Round 3 information for Panel 25, Round 5 information for Panel 24, and Round 7 information for Panel 23, the 2021 PUF has the most up-to-date version of Round 3-Round 5-Round 7 information. These data are the most consistent with subsequent rounds in 2021.

Users who wish to combine data on all rounds for an individual can obtain prior round data from the MEPS HC files noted in the table above.

Following imputation, no values of -2, -7, -8, or -15 remain on any variable. Due to skip patterns, the majority of -1s (question was not asked due to skip pattern) remain. For reasons of confidentiality, values of -10 (hourly wage was top-coded at \$105.77) also remain and employer size (number of employees in establishment) is top-coded at 500.

Variable Naming

The names of the imputed variables are similar to the names of the corresponding constructed variables in this file. An ‘H’ suffix is added and the resulting name is shortened to 8 characters when necessary (e.g., the imputed version of SELFCM31 is SLFCM31H). The variables CMJHLD31/42/53 differ from this naming convention because they are not imputed (they contain no missing values) but were constructed using information from the Person Round Plan (PRPL) File (MEPS HC-232). CMJHLD31/42/53 may be compared with the constructed Employment variables HELD31X/42X/53X. Both of these sets of variables provide information on the insurance coverage individuals obtain through their current main jobs. However, these variables may differ since they are constructed from two different data files and because the CMJHLD31/42/53 variables capture information on some additional sources of employment-related insurance that were identified in the insurance section of the instrument while the HELD31X/42X/53X variables only contain information on sources of coverage identified in the employment section. In addition, for rounds that cross the calendar year from 2020 to 2021 (Panel 23 Round 7, Panel 24 Round 5, and Panel 25 Round 3) the values of CMJHLD31 reflect coverage in the 2021 PRPL file as well as in the 2020 PRPL file. Note that the variables CMJHLD31/42/53 are included on this file because they were used to perform logical edits on the OFFER31H/42H/53H and OFEMP31H/42H/53H variables (edits are described below). The CMJHLD31H/42H/53H variables were used to edit the OFFER31H/42H/53H and OFEMP31H/42H/53H variables (rather than the HELD31X/42X/53X variables) because they were more consistent with the other health insurance variables on the FY PUF for each year.

The following table provides the name of the constructed Employment variables that correspond with each imputed Employment variable.

Employment Variable Crosswalk

Imputed	Constructed
EMPST31H	EMPST31
EMPST42H	EMPST42

Imputed	Constructed
EMPST53H	EMPST53
SLFCM31H	SELFCM31
SLFCM42H	SELFCM42
SLFCM53H	SELFCM53
NMEMP31H	NUMEMP31
NMEMP42H	NUMEMP42
NMEMP53H	NUMEMP53
MORE31H	MORE31
MORE42H	MORE42
MORE53H	MORE53
INDCT31H	INDCAT31
INDCT42H	INDCAT42
INDCT53H	INDCAT53
OCCCT31H	OCCCAT31
OCCCT42H	OCCCAT42
OCCCT53H	OCCCAT53
HOUR31H	HOUR31
HOUR42H	HOUR42
HOUR53H	HOUR53
JBORG31H	JOBORG31
JBORG42H	JOBORG42
JBORG53H	JOBORG53
UNION31H	UNION31
UNION42H	UNION42
UNION53H	UNION53
BSNTY31H	BSNTY31
BSNTY42H	BSNTY42
BSNTY53H	BSNTY53
HRWG31H	HRWG31X
HRWG42H	HRWG42X
HRWG53H	HRWG53X
CMJHLD31*	HELD31X

Imputed	Constructed
CMJHLD42*	HELD42X
CMJHLD53*	HELD53X
OFFER31H	OFFER31X
OFFER42H	OFFER42X
OFFER53H	OFFER53X
OFEMP31H	OFREMP31
OFEMP42H	OFREMP42
OFEMP53H	OFREMP53
PYVAC31H	PAYVAC31
PYVAC42H	PAYVAC42
PYVAC53H	PAYVAC53
SCPAY31H	SICPAY31
SCPAY42H	SICPAY42
SCPAY53H	SICPAY53
PAYDR31H	PAYDR31
PAYDR42H	PAYDR42
PAYDR53H	PAYDR53
RTPLN31H	RETPLN31
RTPLN42H	RETPLN42
RTPLN53H	RETPLN53

* Both CMJHLD and HELDX reflect the insurance status at a current main job but are constructed from different sources of data. See the text for a description of possible differences in these variables. CMJHLD is *not* an imputed version of HELDX.

Imputation Strategy

The first variables to be imputed were the employment status variables (EMPST31H/42H/53H) which identify all persons (EMPST31H/42H/53H = 1 or 2) who should have valid information related to their current main job.

EMPST31H/42H/53H are created from EMPST31/42/53. The EMPST31/42/53 variables have separate response categories for individuals who were “employed during the reference period” and those who were “not employed with no job to return to.” In the imputed variables EMPST31H/42H/53H these responses are collapsed into a single category for analytic purposes. The resulting three responses are mutually exclusive as follows:

Value	Label
-1	INAPPLICABLE
1	EMPLOYED AT RD##/## INT DATE
2	JOB TO RETURN TO AT RD ##/## INT DATE
34	NOT EMPL AT INT DATE/NOT EMPL DUR RD ##/##

Respondents with EMPST31H/42H/53H equal to 1 or 2 go through the imputation process. Respondents with EMPST31H/42H/53H equal to -1 or 34 have values of -1 for all remaining imputed employment variables.

The next set of variables to be imputed were the self-employment/wage-earner variables (SLFCM31H/42H/53H) which determine skip patterns for the remaining variables (e.g., self-employed persons are not asked about wages).

Many of the remaining variables were imputed separately for wage-earners and the self-employed for the following reasons: 1) self-employed and wage-earners were asked different sets of questions about their current main jobs; 2) even when variables were asked for both groups, the quality of the imputations was improved by specifying separate sets of class variables tailored to the wage-earner and self-employed populations.

The weighted sequential hot decking process requires class variables to impute missing values. These class variables were identified with regression models in order to identify the predictive quality of a set of variables for each variable to be imputed. The set of possible class variables includes age, gender, region, educational attainment, industry and occupation code. The same set of variables has been consistently used to impute a given variable across panels.

As mentioned above, self-employed and wage-earners were asked different sets of questions about their current main job. These variables can be categorized into the following sets:

Variables with Valid Data for Wage-Earners but Not for Self-Employed Individuals

- Hourly wage (HRWG31H, HRWG42H, HRWG53H)
- Paid sick leave (SCPAY31H, SCPAY42H, SCPAY53H)
- Paid leave to visit a doctor (PAYDR31H, PAYDR42H, PAYDR53H)
- Paid vacation (PYVAC31H, PYVAC42H, PYVAC53H)
- Pension plan (RTPLN31H, RTPLN42H, RTPLN53H),
- Business has more than one location (MORE31H, MORE42H, MORE53H)
- Sector: private-foreign govt/federal government/state-local government (JBORG31H, JBORG42H, JBORG53H)

Variables with Valid Data for Self-Employed Individuals but Not for Wage-Earners

- Business was incorporated, a proprietorship, or a partnership (BSNTY31H, BSNTY42H, and BSNTY53H)

Variables with Valid Data for All Workers

- Usual hours worked per week (HOUR31H, HOUR42H, HOUR53H)
- Number of employees (NMEMP31H, NMEMP42H, NMEMP53H)
 - Establishment size for wage-earners
 - Business size for self-employed
- Union status (UNION31H, UNION42H, UNION53H)
- Industry category (INDCT31H, INDCT42H, INDCT53H)
- Occupation category (OCCCT31H, OCCCT42H, OCCCT53H)

Variables with Valid Data for All Workers Except for Self-Employed Individuals with No Employees (i.e., persons for whom SLFCM31H/42H/53H = 1 and NMEMP31H/42H/53H = 1 within each round)

- Insurance coverage from current main job (CMJHLD31, CMJHLD42, CMJHLD53)
- Eligible for insurance offered at current main job (OFFER31H, OFFER42H, OFFER53H)
- Insurance offered to anyone at current main job (OFEMP31H, OFEMP42H, OFEMP53H)

In situations where SLFCM31/42/53 is imputed to “Self employed at the current main job” because SELFCM31/42/53 is set to “Refused” (-7) or “Don’t Know” (-8), imputed job characteristic and wage variables are set to “Inapplicable” (-1). Corresponding unimputed variables, however, will contain job characteristic and wage information. This is because persons who indicate “Refused” (-7) or “Don’t Know” (-8) during the interview when asked if the jobholder is self-employed at the current main job follow skip patterns for persons who are not self-employed. Job characteristic and wage information is collected.

Additional Detail on Specific Variables

Hourly Wage (HRWG31H, HRWG42H, HRWG53H)

Valid data for imputed hourly wages is available for all wage-earners (SLFCM31H/42H/53H equal to 2). The values for the imputed hourly wage variables (HRWG31H/42H/53H) reflect the most up-to-date version of hourly wages for the wage-earner in each round. By contrast, the

constructed hourly wage variables (HRWG31X/42X/53X) identify the wage reported in the round a current main job is first reported. Information on any wage changes after that round are contained in the variables DIFFWG31/42/53 and NHRWG31/42/53. These variables, as well as HRWG31X/42X/53X are used in the construction of the imputed hourly wage variables.

For reasons of confidentiality, the hourly wage variables are top-coded. Like the constructed hourly wage variables, imputed hourly wages greater than or equal to \$105.77 are top-coded to a value of -10.

Union Membership (UNION31H/42H/53H)

In addition to using weighted sequential hot-decking techniques, individuals who were identified as being employed by the military had their union membership status logically edited to “Inapplicable” (-1).

Health Insurance (CMJHLD31, CMJHLD42, CMJHLD53, OFFER31H, OFFER42H, OFFER53H, OFEMP31H, OFEMP42H, OFEMP53H)

This file includes several employment-related health insurance variables, CMJHLD31/42/53, OFFER31H/42H/53H, and OFEMP31H/42H/53H. These variables are valid for wage-earners (SLFCM31H/42H/53H = 2) and self-employed individuals with employees (SLFCM31H/42H/53H = 1 and NMEMP31H/42H/53H > 1).

The variables CMJHLD31/42/53 are constructed from the Person-Round-Plan (PRPL) public use file and indicate whether the person held insurance coverage from his/her current main job at some point during the round. This is primarily defined using the PRPL public use file variable CMJINS which identifies insurance obtained through a current main job, and STAT1-12, which indicates whether the coverage is in effect during the month in the delivery year. Since Panel 23 Round 7, Panel 24 Round 5, and Panel 25 Round 3 cross two calendar years, CMJHLD31 indicates if the person held coverage during the portion of Round 7, Round 5, or Round 3 in either the 2020 or 2021 PRPL public use file for the individual.

The variables OFFER31H/42H/53H indicate whether the person was offered (was eligible for) insurance at their current main job at some point during the round. For records that had a value of “Cannot be Computed” (-15), valid reported values for OFFER31H/42H/53H were brought forward from a previous round (including imputed values from that round) if the person did not change jobs before any further imputations were performed.

The variables OFEMP31H/42H/53H indicate whether an employer offered health insurance to any employees in the establishment and rely on OFFER31H/42H/53H during their construction.

OFFER31H/42H/53H and the related variable OFEMP31H/42H/53H were logically edited as follows. If a person indicates that they held insurance from their current main job (CMJHLD31/42/53 = 1) then OFFER31H/42H/53H was set equal to 1. For rounds that cross the calendar year from 2020 to 2021 (Panel 23 Round 7, Panel 24 Round 5, Panel 25 Round 3), OFFER31H/53H was set to “Yes” (1) if the person held coverage at any point in the round in either the first or second calendar year of the round. If a person indicated that they held insurance (CMJHLD31/42/53 = 1) or were offered insurance (OFFER31H/42H/53H = 1) at their current

main job, then OFEMP31H/42H/53H, the variable indicating that the employer offered insurance to at least one employee, was set equal to “Yes” (1).

Number of Employees (NMEMP31H, NMEMP42H, NMEMP53H)

Like the corresponding constructed NUMEMP31/42/53, NMEMP31H/42H/53H indicate the number of employees at the location of the person’s current main job. Due to confidentiality concerns, this variable has been top-coded at 500. Missing value imputation is done using weighted sequential hot-decking techniques. Note that the definition of NMEMP31H/42H/53H, like that for NUMEMP31/42/53, differs for wage-earners and self-employed individuals. For wage-earners, it represents the size of the worker’s establishment. For self-employed individuals, it represents the size of the self-employed individual’s entire business.

2.5.10 Health Insurance Variables (TRIJAyyX-PMEDPY53)

Throughout Section 2.5.10 references to yy represent the year, 21, references to mm indicate the month (JA through DE), and references to rr indicate a combination of rounds (31/42/53, where the first r denotes the interview round for Panel 25 and the second r denotes the round for Panel 26) or the end of the calendar year (21). For the two extended panels - Panel 23 and Panel 24 - the “31” variables contain data from Round 7 (Panel 23) or Round 5 (Panel 24), the “42” variables contain data from Round 8 (Panel 23) or Round 6 (Panel 24), and the “53” variables contain data from Round 9 (Panel 23) or Round 7 (Panel 24).

Beginning Panel 22 Round 3/Panel 23 Round 1, design changes to the health insurance section may impact trend analyses. Analysts should note that a series of questions were added to the HX section of the questionnaire to confirm whether a person who did not initially report any comprehensive coverage during a round has insurance. Starting at HX210, questions were presented to respondents who at that point in the instrument had not yet reported any sources of health insurance coverage, or only reported a source of health insurance without hospital and physician benefits, to determine whether they had coverage that included hospital and physician benefits. If the respondent answered affirmatively at HX210, subsequent questions identified the specific type of coverage (e.g. Medicaid, Private, etc.). This may cause analysts to see changes to the insurance variables-particularly, changes to the monthly health insurance coverage indicators: PUBmmyyX, PRImmyyX, INSmmyyX; and the summary health insurance coverage indicators: UNINSyy, INSCOVyy, INSURCYy, PUBrrX, PUBATrrX, PRIVrr, PRIVATrr, INSrrX, and INSATrrX.

Other changes were made in FY 2018 to the health insurance questions that may affect the continuity of estimates. These changes include modifications to the Medicaid/SCHIP, and TRICARE/CHAMPVA questions to ask if each person in the household is covered using the person’s name in the question text (e.g. “Was Person 1 covered?” “What about Person 2?” etc.). Additionally, in Rounds 2 and 3, respondents are now required to answer “Yes” or “No” for each person individually when reviewing coverage from a previous round for these insurance sources. Changes to the Medicare Round 1 series were also made to probe separately for persons in the RU who were 65 years of age or older versus RU members who were under 65 years of age. Similar to the Medicaid and TRICARE series, Medicare coverage questions were asked for each RU member who was at least 65 years old.

The aforementioned changes to the administration of the insurance section may also be evident in the Managed Care Variables (TRISTyyX-PRVHMOyy) because more respondents are now more likely to be asked about managed care.

Respondents were allowed to simultaneously report Medicaid and other public hospital/physician coverage. Analysts should be aware that they might see changes in coverage trends in the constructed variables relating to Medicaid, edited Medicaid, or Other Public coverage as well as respondents reporting both after FY 2018.

The variables VERFLG31, VERFLG42, and VERFLGyy indicate the round in which comprehensive health insurance coverage was first reported through the verification series of questions collected in the loop that starts at HX210 (HXLoop_40). These values will be carried through to subsequent rounds (e.g., from VERFLG31 to VERFLG42) if the coverage initially added through the verification loop continues, and no other comprehensive source of coverage is reported for that person outside of the verification loop. If previously reported coverage through the verification series ends and, in a future round, new comprehensive coverage is reported through the verification loop, then the VERFLG31/42/yy variable will reflect the corresponding round of the newly reported coverage. The VERFLG variables were set to '95' to indicate that: 1) coverage was reported outside verification; 2) the person did not have coverage; or 3) the person would have been assigned edited coverage even though they may have reported coverage in the verification loop. As an example of the latter, a person who is age 65 or older and reports Medicare coverage through verification but also reports receipt of social security would have MCAREX set to '1' because of the report of social security so the report of coverage in the verification module would not have changed their coverage status in the MEPS. In FY 2019, the construction of the VERFLG variables was modified such that all persons ages 65 and older who gained edited Medicare through the Medicare coverage of their spouse also have a value of 95 in the verification variables, provided that the coverage of the spouse was added outside of the verification series.

Persons who report coverage under Indian Health Service (IHS) are identified in the constructed variables IHSrr, IHSATrr, and IHSmmyy. Persons reporting only IHS coverage are not considered covered for the summary insurance measures PUBmmyyX, PUByyX, INSmmyyX, INSCOVyy, and INSURCYyy. Persons who report coverage under Veteran's Administration (VA) can be identified in this file in the constructed variables VAPROGrr, VAPRATrr, VAEVyy, as well as the monthly variables VAPRmmyy.

Monthly Health Insurance Indicators (TRIJAYyX-INSDEyyX)

Constructed and edited variables are provided that indicate any coverage in each month of 2021 for the sources of health insurance coverage collected during the MEPS interviews (Panel 23 Rounds 7 through 9, Panel 24 Rounds 5 through 7, Panel 25 Rounds 3 through 5, and Panel 26 Rounds 1 through 3). One edit to the private insurance variables corrects for a problem concerning covered benefits that occurred when respondents reported a change in any of their private health insurance plan names. Additional edits address issues of missing data on the time period of coverage for both public and private coverage that was either reviewed or initially reported in a given round. Other edits, described below, were performed on the Medicare and Medicaid or State Children's Health Insurance Program (SCHIP) variables to assign persons to

coverage from these sources. Observations that contain edits assigning persons to Medicare or Medicaid/SCHIP coverage can be identified by comparing the edited and unedited versions of the Medicare and Medicaid/SCHIP variables. Starting October 1, 2001, persons 65 years and older have been able to retain TRICARE coverage in addition to Medicare. Therefore, unlike in earlier MEPS public use files, persons 65 years and older do not have their reported TRICARE coverage (TRIJAYYX - TRIDEYYX) overturned. TRICARE acts as a supplemental insurance for Medicare, similar to Medigap insurance.

Public sources include Medicare, TRICARE/CHAMPVA, Medicaid, SCHIP, and other public hospital/physician coverage. IHS is not included as a public source of coverage.

Medicare

Medicare (MCRJAYY - MCRDEYY) coverage was edited (MCRJAYYX - MCRDEYYX) for persons age 65 or over. Within this age group, individuals were assigned Medicare coverage if:

- They answered “Yes” to a follow-up question on whether they received Social Security benefits; or
- They were covered by Medicaid/SCHIP, other public hospital/physician coverage or Medigap coverage; or
- Their spouse was age 65 or over and covered by Medicare; or
- They reported TRICARE coverage.

Note that age (AGERX) is checked for edited Medicare, however date of birth is not considered. Edited Medicare is somewhat imprecise with regard to a person’s 65th birthday.

Medicaid/SCHIP and Other Public Hospital/Physician Coverage

Questions about other public hospital/physician coverage were asked in an attempt to identify Medicaid or SCHIP recipients who may not have recognized their coverage as such. Beginning Panel 22 Round 3/Panel 23 Round 1, these questions were asked even if a respondent reported Medicaid or SCHIP directly. (Previously, other public hospital/physician coverage was only asked for respondents who did not report Medicaid or SCHIP.) Respondents reporting other public hospital/physician coverage were asked follow-up questions to determine if the coverage was through a specific Medicaid HMO or if it included some other managed care characteristics. Respondents who identified managed care from either source were asked if the recipient paid anything for the coverage and/or if a government source paid for the coverage.

The Medicaid/SCHIP variables (MCDJAYY - MCDDEYY) have been edited (MCDJAYYX - MCDDEYYX) to include persons who paid nothing for their other public hospital/physician insurance when such coverage was through a Medicaid HMO or reported to include some other managed care characteristics.

To assist users in further editing sources of insurance, this file contains variables constructed from the other public hospital/physician series that indicate:

- Respondents who reported coverage through other public hospital/physician insurance (GVAJAyy - GVADEyy);
- Respondents who reported coverage through other public hospital/physician insurance and the source was identified as an HMO or the source of coverage required the covered person to sign up with a doctor or clinic (GVBJAyy - GVBDEyy);
- Respondents who reported coverage through other public hospital/physician insurance and the person pays a monthly premium for this coverage (GVCJAyy - GVCDEyy)

The variables GVAJAyy - GVADEyy, GVBJAyy - GVBDEyy, and GVCJAyy - GVCDEyy are provided only to assist in editing and should not be used to make separate insurance estimates for these types of insurance categories.

Any Public Insurance in Month

The file also includes summary measures that indicate whether or not a sample person has any public insurance in a month (PUBJAyyX - PUBDEyyX). Persons identified as covered by public insurance are those reporting coverage under TRICARE, Medicare, Medicaid or SCHIP, other public hospital/physician programs, or Veteran's Administration (VA). IHS is not included as a public source of coverage. Note that further edits may be made to the public insurance variables in later MEPS data releases to address cases where private coverage through a federally-facilitated, state-based or state partnership exchange/marketplace may have been originally reported as public insurance. These potential edits could affect the variables MCAIDyyX, GOVTAAyy, GOVTByy, GOVTCyy, and PUByyX.

Private Insurance

Variables identifying private insurance in general (PRIJAyy - PRIDEyy) and specific private insurance sources [such as employer/union group insurance (PEGJAyy - PEGDEyy); non-group (PNGJAyy - PNGDEyy); other group (POGJAyy - POGDEyy)]; and private insurance through a federally-facilitated, state-based or state partnership exchange/marketplace (PRXJAyy - PRXDEyy) were constructed. Private insurance sources identify coverage in effect at any time during each month of 2021. Separate variables beginning with the letter "H" identify policyholders (e.g., HPEJAyy - HPEDEyy). Both types of variables indicate coverage or policyholder status for a particular source and do not identify persons who may have more than one policy of a given source where they are just covered or are also a policyholder (for example, someone who is a policyholder for one employer/union group plan and also a dependent on another employer/union group plan held by his/her spouse). In some cases, the policyholder was unable to characterize the source of insurance (PDKJAyy - PDKDEyy).

Prior to FY 2018, persons covered under policyholders living outside the RU were identified in POUJAyy - POUDEyy and PROUTr. Beginning FY 2018, the constructed variables PRIEUOr and PRINEOr are included. PRIEUOr indicates coverage from a policyholder living outside the

RU where the source is through an employer, and PRINEOrr indicates coverage from a policyholder living outside the RU where the source is not through an employer. These variables are based on responses to a follow-up question for respondents who indicate coverage from a policyholder outside the household. The question HP130 asks “Is the {INSURANCE SOURCE NAME} health coverage {POLICYHOLDER} has through an employer or previous employer?” If the respondent’s answer to HP130 was unknown, the person’s coverage is now included in PRIDKrr.

An individual was considered to have private health insurance coverage if, at a minimum, that coverage provided benefits for hospital and physician services (including Medicare supplemental coverage). Note, however, that persons covered by private insurance through an exchange/marketplace (PRSTXrr and PRXJAyy - PRXDEyy) were considered to have private health coverage if that coverage provided hospital/physician services, but excluded coverage that was explicitly identified as Medicare supplemental coverage (HX620/OE130=5). If a person reported Medicare supplemental coverage through the exchange/marketplace, then the source of the insurance purchased was edited to reflect coverage “from a professional association” (HP40=1) or coverage “from a group or association” (HX200/HX300=1). Further descriptions of the exchange variables are detailed below. Sources of insurance with missing information regarding the type of coverage were assumed to contain hospital/physician coverage. Persons who reported private insurance that did not provide hospital/physician insurance were not counted as privately insured. Coverage indicated by these variables may be from any type of job whereas the employment section insurance variables delivered on this file reflect only coverage through a current main job.

Health insurance through a job or union (PEGJAyy - PEGDEyy) was initially asked about in the Employment Section of the interview and later confirmed in the Health Insurance Section.

Insurance that was reported in the employment section through a job classified as self-employed with firm size of 1 is included in the other private insurance variables: PEGJAyy - PEGDEyy; PNGJAyy - PNGDEyy; POGJAyy - POGDEyy; PDKJAyy-PDKDEyy; HPEJAyy-HPEDDEyy; HPNJAyy-HPNDEyy; HPOJAyy-HPODEyy; HPDJAyy-HPDDEyy; and PRIEUrr, PRINGrr, PRIOGrr, and PRIDKrr based on responses at HP40.

Private insurance that was not employment-related (POGJAyy - POGDEyy, PNGJAyy - PNGDEyy, PDKJAyy - PDKDEyy, PNEJAyy - PNEDEyy, and PRXJAyy - PRXDEyy) was reported in the Health Insurance section only.

Beginning in Panel 14 Round 5/Panel 15 Round 3, “High Risk Pool” was added to the list of categories (HX03 =10 and HX23 =13). Beginning FY 2010, High Risk Pool was included in all Other Group insurance categories. Beginning in Panel 22 Round 3/Panel 23 Round 1, the response category “High Risk Pool” was removed from HP40, HX200, and HX300.

“Federal/State Exchange” is included in the list of private insurance categories (HP40=8 and HX200/HX300 =11). Information on federal/state exchanges is also collected at question HP50 (“Is this coverage through {state exchange name}?”) for respondents reporting insurance from a group, directly from an insurance company or HMO, from an insurance agent or from an “other” unspecified source and at OE40 in Rounds 3, 5, and 7 only (“Is this coverage through {state

exchange name}?”) for respondents who previously reported private insurance coverage from an insurance company or HMO, or from an insurance agent that was not through an exchange/marketplace. Note that the state-specific name for the exchange/marketplace was used when asking these questions and was also used on the list of private insurance categories at HP40, HX200, and HX300. The variables PRSTXrr have been constructed to include persons less than 65 years old who report private insurance through a federally-facilitated, state-based or state partnership exchange/marketplace at HP40, HX200, or HX300, or persons 65 years old or older who report private insurance through a federally-facilitated, state-based or state partnership exchange/marketplace at HP40, HX200, or HX300 and who were not covered by Medicare. In addition, persons who reported a source of insurance at HX200 or HX300 that was not through an exchange/marketplace (e.g., through a group or directly from an insurance company) but who answered yes to HP50 or OE40 were also classified as having exchange/marketplace coverage instead of being assigned to the category they originally reported. In addition to reporting coverage through an exchange/marketplace, coverage needed to have been identified as hospital/physician coverage at HX620/OE130 (=1 or missing (-7, -8)), but not identified as having Medicare supplemental coverage (HX620/OE130=5). The variables PRSTXrr contain information on private coverage that was reported as obtained through a federally-facilitated, state-based or state partnership marketplace. Consistent with the approach used in the Current Population Survey and the National Health Interview Survey, MEPS respondents reporting public coverage were asked whether the public coverage was obtained through a federal or state marketplace in case respondents were confused about whether the source of coverage was public or private. Responses to these questions were not used to edit the PRSTXrr variables.

Any Insurance in Month

The file also includes summary measures that indicate whether or not a person has any insurance in a month (INSJAyyX - INSDEyyX). Persons identified as insured are those reporting coverage under TRICARE, Medicare, Medicaid, SCHIP, other public hospital/physician or private hospital/physician insurance (including Medigap plans), or Veteran’s Administration (VA). A person is considered uninsured if not covered by one of these insurance sources. IHS is not included as a source of coverage.

Summary Insurance Coverage Indicators (PRVEVyy-INSURCyy)

The variables PRVEVyy-UNINSyy summarize health insurance coverage for the person in 2021 for the following types of insurance: private (PRVEVyy); TRICARE/CHAMPVA (TRIEVyy); Medicaid or SCHIP (MCDEVyy); Medicare (MCREVyy); other public coverage (GVAEVyy); other public coverage that is an HMO (GVBEVyy); other public coverage where a premium is paid (GVCEVyy). Each variable was constructed based on the values of the corresponding 12 month-by-month health insurance variables described above. For persons not in scope for the full year, these summary variables are based on the period of eligibility. If the person was not in scope for all 12 months throughout the year, the values are based on the months the person was eligible. A value of 1 indicates that the person was covered for at least one day of at least one month during 2021. A value of 2 indicates that the person was not covered for a given type of insurance for all of 2021. The variable UNINSyy summarizes PRVEVyy-GVAEVyy. Where PRVEVyy-GVAEVyy are all equal to 2, then UNINSyy equals 1, person was uninsured for all of 2021. Otherwise, UNINSyy is set to 2, insured for all or part of 2021.

For user convenience, this file contains a constructed variable INSCOVyy that summarizes health insurance coverage for the person in 2021, with the following three values:

1 = ANY PRIVATE (Person had any private insurance coverage [including TRICARE/CHAMPVA] any time during 2021)

2 = PUBLIC ONLY (Person had only public insurance coverage [excluding TRICARE/CHAMPVA] during 2021)

3 = UNINSURED (Person was uninsured during all of 2021)

INSURCy summarizes health insurance coverage for the person in 2021 using eight categories of insurance separated by age using the person's age on December 31st, 2021:

1 = ANY PRIVATE (0-64) (Person is between 0 and 64 years old and is covered by private insurance or TRICARE/CHAMPVA in 2021)

2 = PUBLIC ONLY (0-64) (Person is between 0 and 64 years old and is covered by public insurance only (excluding TRICARE/CHAMPVA) in 2021)

3 = UNINSURED (0-64) (Person is between 0 and 64 years old and is uninsured for all of 2021)

4 = EDITED MEDICARE ONLY (65+) (Person is 65 years old or more and is covered by edited Medicare only in 2021)

5 = EDITED MEDICARE & PRIV (65+) (Person is 65 years old or more and is covered by edited Medicare and private insurance or TRICARE/CHAMPVA in 2021)

6 = EDITED MEDICARE & OTH PUB ONLY (65+) (Person is 65 years old or more and is covered by edited Medicare and public insurance including edited Medicaid/SCHIP or other public coverage but excluding TRICARE/CHAMPVA in 2021)

7 = UNINSURED (65+) (Person is 65 years old or more and is uninsured for all of 2021)

8 = NO MEDICARE BUT ANY PUBLIC/PRIVATE (65+) (Person is 65 years old or more and is not covered by Medicare but is covered by private insurance, Medicaid, TRICARE/CHAMPVA, Veteran's Administration, or other public coverage in 2021)

Please note, beginning in 2012, Category 7 was revised to categorize persons who are 65 years or older and uninsured, and Category 8 was added to include persons 65 years or older who do not have Medicare, but are covered by public or private insurance.

Please note that IHS is not included as a source of coverage for either INSCOVyy or INSURCy.

Please note that both INSCOVyy and INSURCy categorize TRICARE as private coverage. All other health insurance indicators included in this data release categorize TRICARE as public

coverage. If an analyst wishes to consider TRICARE public coverage, the variable can easily be reconstructed using the PRVEVyy and TRIEVyy variables. Also note that these categories are mutually exclusive, with preference given to private insurance and TRICARE. Persons with both private insurance/TRICARE and public insurance will be coded as “1” for INSCOVyy and INSURCyy.

Users wishing to compare INSCOVyy and INSURCyy across years should note at least two changes beginning in 2018 that may affect the continuity of estimates: 1) increased reports of coverage due to the inclusion of the coverage verification series; and 2) the inclusion of Veteran’s Administration coverage as a public coverage source.

Flexible Spending Accounts (FSAGT31-PFSAMT31)

Respondents in Rounds 1, 3, 5, and 7 were asked if any RU members set aside pre-tax dollars of their own money to pay for out-of-pocket health care expenses. If an RU has a Flexible Spending Account (FSA), then FSAGT31 was set to 1 (Yes), and two follow-up questions were asked - HASFSA31 and PFSAMT31. HASFSA31 was set for each RU member to indicate which RU member has an FSA. The constructed variable PFSAMT31 indicates the total amount the individual RU member contributed to their FSA. If no RU member has an FSA, then both HASFSA31 and PFSAMT31 are set to -1 (Inapplicable).

Unedited Health Insurance Variables (PREVCOVR-MORECOVR)

Duration of Uninsurance

If a person was identified as being without insurance as of January 1st in the MEPS Round 1 interview, a series of follow-up questions was asked to determine the duration of uninsurance prior to the start of the MEPS survey. Persons who were insured as of January 1st, and persons with a date of birth on or after December 31, 2021 or whose age category was less than 1 year old were skipped past this loop of questions. These questions are asked in Round 1 only.

PREVCOVR indicates if the person was covered by insurance in the two years prior to the MEPS Round 1 interview. For persons who reported only non-comprehensive coverage as of January 1st, a question was asked to determine if they had been covered by more comprehensive coverage that paid for medical and doctors’ bills in the previous two years (MORECOVR). Beginning Panel 23 Round 1, the follow-up questions to PREVCOVR and MORECOVR that collected information on the most recent month and year of coverage (COVRMM, COVRY, INSENDMM, INSENDYY) and type of coverage (Employer-sponsored (WASESTB), Medicare (WASMCARE), Medicaid/SCHIP (WASMCAID), TRICARE/CHAMPVA (WASCHAMP), VA/Military Care (WASVA), Other public (WASOTGOV, WASAFDC, WASSSI, WASSTAT1-4, WASOTHER), private coverage purchased through a group, association or insurance company (WASPRIV)) are no longer asked. Therefore, these variables will no longer be constructed.

Note that these variables are unedited and have been taken directly as they were recorded from the raw data. There may be inconsistencies with the health insurance variables released on public use files that indicate that an individual is uninsured in January. Out-of-scope persons in both panels have been set to “Inapplicable” (-1) for PREVCOVR and MORECOVR. All other

persons have PREVCOVR and MORECOVR copied directly from the value of the unedited source variable.

Persons whose January 1st insurance coverage status could not be determined due to their reference period beginning after January 1st were also asked the follow-up questions described above. In these cases, persons who reported comprehensive coverage were asked if they were ever without insurance. Those who were uninsured were asked to determine the duration of uninsurance prior to the start of their reference period. Those who reported only non-comprehensive coverage were asked if they had been covered by comprehensive coverage that paid for medical and doctors' bills in the previous two years. Coverage is determined by health insurance status during the whole reference period or the month of January and ignores that these persons were not in the household on January 1st.

Health Insurance Coverage Variables - At Any Time/At Interview Date/At 12-31 Variables (TRICR31X-INSATyyX)

Constructed and edited variables are provided that indicate health insurance coverage at any time in a given round, as well as at the MEPS interview dates and on December 31, 2021. Note that for persons who left the RU before the MEPS interview date or before December 31st, the variables measuring coverage at the interview date or on December 31st represent coverage at the date the person left the RU.

Variables indicating coverage for Panel 25 members for any time in the round that end in "31" indicate coverage for the portion of Round 3 that occurred in calendar year 2021, unless noted otherwise (see "Dental and Prescription Drug Private Insurance" section). This is also true for data from Panel 24 Round 5 and Panel 23 Round 7 - the 2021 portion of those rounds is contained in the "31" variables. Variables indicating coverage for Panel 26 members ending in "53" indicate coverage at any time in Round 3, including the portion of the round that occurred in calendar year 2022. Similarly, all of the Panel 24 Round 7 data, including data collected in 2022, are contained in the "53" variables. For Round 3 coverage for Panel 26 members or Round 7 data for Panel 24 members that occurred in calendar year 2021, users should use variables ending in "yy". The Panel 23 Round 8 data, Panel 24 Round 6 data, Panel 25 Round 4 data, and Panel 26 Round 2 data are stored in the 42 variables.

The health insurance variables are constructed for the sources of health insurance coverage collected during the MEPS interviews (Panel 23 Rounds 7 through 9, Panel 24 Rounds 5 through 7, Panel 25 Rounds 3 through 5, and Panel 26 Rounds 1 through 3). Note that the Medicare variables on this file as well as the private insurance variables that indicate the particular source of private coverage (rather than any private coverage) only measure coverage at the interview date and on December 31st. Users should also note that the same general editing rules were followed for the month-by-month health insurance variables released on this public use file (see the section "Monthly Health Insurance Indicators" for details). Editing programs checking for consistencies between these sets of variables were developed in order to provide as much consistency as possible between the round-specific indicators and the month-by-month indicators of insurance.

Public sources include Medicare, TRICARE, Veteran’s Administration (VA), Medicaid/SCHIP, and other public hospital/physician coverage. IHS was not considered a public coverage source.

Medicare

Medicare coverage variables (MCARErr) and the edited versions of these variables (MCARErrX) were constructed similarly to the month-by-month Medicare variables. Since Medicare coverage is logically edited to continue for a person once it has been reported in MEPS, the Medicare coverage variables can be considered as either “coverage at any time in the round” or “coverage at the interview date” variables. The same caveats as noted above regarding persons who left the RU prior to the interview date or regarding coverage on December 31st variables and restrictions on Round 3, Round 5, and Round 7 coverage to reflect coverage in 2021.

Medicaid/SCHIP and Other Public Hospital/Physician Coverage

Medicaid/SCHIP variables (MCAIDrr) and the edited versions of these variables (MCAIDrrX and MCDATrrX) were constructed similarly to the month-by-month Medicaid/SCHIP variables.

Other public coverage variables indicating coverage through other public hospital/physician insurance (GOVTarr and GOVAATrr); other public coverage that is an HMO (GOVTBrr and GOVBATrr); and other public coverage that requires a premium (GOVTCrr and GOVCATrr) were constructed similarly to the month-by-month Other Public variables.

Any Public Insurance

Any public insurance variables (PUBrrX and PUBATrrX) were constructed similarly to the month-by-month any public insurance variables. The variables indicating coverage through Veteran’s Administration (VAPROGrr and VAPRATrr) are included in this file and were constructed similarly to the Veteran’s Administration month-by-month variables.

Private Insurance

Variables identifying private insurance in general (PRIVrr and PRIVATrr) and specific private insurance sources (such as employer/union group insurance [PRIEUrr]; other group coverage [PRIOGrr]; coverage through an unknown private category [PRIDKrr]; coverage from a policyholder living outside the RU that is employer based coverage [PRIEUOrr]; coverage from a policyholder living outside the RU that is not employer-based coverage [PRINEOrr]; and coverage through an exchange [PRSTXrr]) were constructed similarly to the month-by-month variables in the section “Monthly Health Insurance Indicators.” Variables indicating any private insurance coverage are available for the following time periods: at any time in a given round, at the interview date, and on December 31st. The variables for the specific sources of private coverage are only available for coverage on the interview dates and on December 31st.

Any Insurance in Period

Any insurance variables (INSrrX and INSATrrX) were constructed similarly to the month-by-month any insurance program variables.

FY 2021 PUF Managed Care Variables (TRIST31X- PRVHMOyy)

In addition to the month-by-month indicators of coverage, there are round-specific health insurance variables indicating coverage by an HMO or managed care plan. Managed care

variables have been constructed from information on health insurance coverage at any time in a reference period and the characteristics of the plan. A separate set of managed care variables has been constructed for private insurance, Medicaid/SCHIP, and Medicare coverage. The purpose of these variables is to provide information on managed care participation during the portion of the three rounds (i.e., reference periods) that fall within the same calendar year.

Managed care variables for calendar year 2021 are based on responses to health insurance questions asked during the Round 7, 8, and 9 interviews of Panel 23, the Round 5, 6, and 7 interviews of Panel 24, the Round 3, 4, and 5 interviews of Panel 25, and the Round 1, 2, and 3 interviews of Panel 26. Each managed care variable ends in “rr” where the first r denotes the interview round for Panel 25 and the second r denotes the round for Panel 26, respectively. For the two extended panels - Panel 23 and Panel 24 - the “31” variables contain data from Round 7 (Panel 23) or Round 5 (Panel 24), the “42” variables contain data from Round 8 (Panel 23) or Round 6 (Panel 24), and the “53” variables contain data from Round 9 (Panel 23) or Round 7 (Panel 24). The variables ending in “31” and “42” correspond to the first two interviews of each panel in the calendar year. Because Round 3 interviews typically overlap the final months of one year and the beginning months of the next year, the “31” managed care variables for Panels 24 and 25 indicate whether or not a person has coverage from a managed care plan in the 2021 calendar year. This is also the case for the Panel 23 Round 7 and Panel 24 Round 5 data - the “31” managed care variables are limited to the 2021 calendar year. Similarly, the Panel 24 Round 7 and Panel 26 Round 3 managed care variables indicate whether or not a person has coverage from a managed care plan in the 2021 calendar year, and the variables have been given the suffix “yy” (as opposed to “53”) to emphasize the restricted time frame. Further descriptions of the implications to managed care plan coverage due to the overlapping calendar year in Rounds 3 and 5 are detailed below.

Construction of the managed care variables is straightforward, but three caveats are appropriate. First, MEPS estimates of the number of persons in HMOs are higher than figures reported by other sources, particularly those based on HMO industry data. The differences stem from the use of household-reported information, which may include respondent error, to determine HMO coverage in MEPS.

Second, the managed care questions are asked about the last plan held by a person through their establishment (employer or insurer) even though the person could have had a different plan through the establishment at an earlier point during the interview period. As a result, in instances where a person changed their establishment-related insurance, the managed care variables describe the characteristics of the last plan held through the establishment.

Third, the “yy” versions of the managed care variables are developed from two sets of Rounds 3, 5, and 7 source variables that cover different time frames. Using Round 3 managed care variables as an example, the first set of source variables - Round 3 health insurance status variables - are restricted to the same calendar year as the Round 1 and Round 2 data. The second set of source variables - Round 3 variables describing plan type - overlap the next calendar year, 2022. As a consequence, the “yy” managed care variables may not describe the characteristics of the last plan held in the calendar year if the person changed plans in the beginning of the following year.

The variables PRVHMOr indicate coverage by a private HMO in Panel 26 Rounds 1 - 3, Panel 25 Rounds 3 - 5, Panel 24 Rounds 5 - 7, and Panel 23 Rounds 7 - 9. The variables MCRPHOr indicate coverage by a Medicare managed care plan (or “Medicare Advantage” plan) in Panel 26 Rounds 1 - 3, Panel 25 Rounds 3 - 5, Panel 24 Rounds 5 - 7, and Panel 23 Rounds 7 - 9. The variables MCRPDrr indicate coverage by Medicare prescription drug benefit, also known as Part D, in Panel 26 Rounds 1 - 3, Panel 25 Rounds 3 - 5, Panel 24 Rounds 5 - 7, and Panel 23 Rounds 7 - 9. The edited version of the Medicare prescription drug coverage variables (MCRPDrrX) include persons who are covered by both edited Medicare and edited Medicaid. The variables MCDHMOr and MCDMCrr indicate coverage by a Medicaid or SCHIP HMO or managed care plan in Panel 26 Rounds 1 - 3, Panel 25 Rounds 3 - 5, Panel 24 Rounds 5 - 7, and Panel 23 Rounds 7 - 9. The Tricare plan variables are similarly defined.

For Panel 26, the “31” version indicates coverage at any time in Round 1, the “42” version indicates coverage at any time in Round 2, and the “yy” version represents coverage at any time during the 2021 portion of Round 3. For Panel 25, the “31” version indicates coverage at any time during the 2021 portion of Round 3, the “42” version indicates coverage at any time in Round 4, and the “yy” version represents coverage at any time during Round 5 since Round 5 ends on 12/31/2021 for Panel 25. For Panel 24, the “31” version indicates coverage at any time during the 2021 portion of Round 5, the “42” version indicates coverage at any time in Round 6, and the “yy” version represents coverage at any time during the 2021 portion of Round 7. For Panel 23, the “31” version indicates coverage at any time during the 2021 portion of Round 7, the “42” version indicates coverage at any time during Round 8, and the “yy” version represents coverage at any time during Round 9, since Round 9 ends on 12/31/2021 for Panel 23.

In the health insurance section of the questionnaire, respondents reporting private health insurance were asked to identify what types of coverage a person had via a checklist. If the respondent selected prescription drug or dental coverage from this checklist, variables were constructed to indicate prescription drug or dental coverage respectively. It should be noted, however, that in some cases respondents may have failed to identify prescription drug or dental coverage that was included as part of a hospital and physician plan.

TRICARE Plan Variables

In Fall 2022 the response options for the CAPI TRICARE questions HX125_01, HX260 and PR280_01 were changed. Options “Tricare Standard”, “Tricare Prime”, “Tricare Extra”, and “Tricare For Life” were replaced by the single response option “Tricare”. As a result the previous plan-specific variables TRICARE Standard (TRISTrrX), TRICARE Prime (TRIPRrrX), TRICARE Extra (TRIEXrrX), and TRICARE for Life (TRILrrX) were dropped from the 2021 data, and new variable TRrrX (Person Covered by Tricare at Any Time during the Reference Period) was added.

Beginning in Panel 9 Rounds 4 and 5/Panel 10 Rounds 1 through 3, CHAMPVA was added to the list of TRICARE/CHAMPVA Plans collected in the instrument. Therefore, the variables TRICH42/yyX were created. The “31” version of this variable was constructed starting in 2006. It should be noted that the TRICARE Plan information was elicited from a pick-list, code-all-that-apply question that asked which type of TRICARE plan the person obtained. Beginning Panel 22 Round 3/Panel 23 Round 1, questions related to military health coverage were asked at the person-level. If it was reported that someone in the RU had coverage through military health

care, a follow-up question was asked to determine who in the RU was covered; then, the pick-list, code-all-that-apply question described above was asked to determine which type of military coverage the person obtained. VA (Veteran's Administration) was added to this list beginning Panel 22 Round 3/Panel 23 Round 1.

In each round, the TRICARE variable has four possible values:

- 1 The person was covered by a TRICARE plan.
- 2 The person was covered by CHAMPVA but not TRICARE.
- 3 The person was not covered by TRICARE/CHAMPVA.
- 1 The person was out-of-scope.

Medicare Managed Care Plans, Part B, and Prescription Drug Benefit

Persons were assigned Medicare coverage based on their responses to the health insurance questions or through logical editing of the survey data. A small number of persons were edited to have Medicare. For this group, coverage through a managed care plan, Part B, and coverage by prescription drug plan questions were not asked. Since no Medicare establishment-person pair exists for this group, the persons' Medicare managed care, Part B, and prescription drug benefit statuses are set to -15 (Cannot be Computed). For those persons who reported Medicare coverage based on their responses to the health insurance questions, the Medicare managed care plan, Part B, and prescription drug benefit questions were asked. Medicare managed care plan and prescription drug benefit questions were asked for each round a person indicates Medicare coverage. Medicare Part B questions were asked during the first report of Medicare only. The Medicare Part B indicator for those persons who indicated not having a Medicare card available was introduced for Panel 14 Round 2 and Panel 13 Round 4. For those persons who reported having Medicare coverage in Round 1, but did not have a Medicare card available, Medicare Part B coverage was set to -15 (Cannot be Computed).

The Medicare prescription drug benefit variables (MCRPDrr) have been edited (MCRPDrrX) to turn on coverage for all persons who are covered by both edited Medicare and edited Medicaid regardless of the status on their unedited Medicare prescription drug benefit variable.

In each round, the variables MCRPHOrr have five possible values:

- 1 The person was covered by Medicare and covered through a Medicare Managed Care or Medicare Advantage Plan.
- 2 The person was covered by Medicare but not covered through a Medicare Managed Care or Medicare Advantage Plan.
- 3 The person was not covered by Medicare.
- 15 The person was covered by Medicare but whether the coverage is through a Medicare Managed Care or Medicare Advantage Plan cannot be computed.

-1 The person was out-of-scope.

In each round, the variables MCRPDrr/MCRPDrrX have five possible values:

1 The person was covered by Medicare and covered by prescription drug benefit.

2 The person was covered by Medicare but not covered by prescription drug benefit.

3 The person was not covered by Medicare.

-15 The person was covered by Medicare but prescription drug benefit coverage cannot be computed.

-1 The person was out-of-scope.

In each round, the variables MCRPBrr have five possible values:

1 The person was covered by Medicare and covered by Part B.

2 The person was covered by Medicare but not covered by Part B.

3 The person was not covered by Medicare.

-15 The person was covered by Medicare but Part B cannot be computed.

-1 The person was out-of-scope.

Medicaid/SCHIP Managed Care Plans

Persons were assigned Medicaid or SCHIP coverage based on their responses to the health insurance questions or through logical editing of the survey data. The number of persons who were edited to have Medicaid or SCHIP coverage is small. These persons indicated coverage through an Other Government program that was identified as being in a Medicaid HMO or gatekeeper plan that did not require premium payment from the insured party. By definition, respondents were asked about the managed care characteristics of this insurance coverage.

Medicaid/SCHIP HMOs

If Medicaid/SCHIP or Other Government programs were identified as the source of hospital/physician insurance coverage, the respondent was asked about the characteristics of the plan. The variables MCDHMOrr were set to “Yes” if an affirmative response was provided to the following question:

Under {Medicaid}, also known as {STATE NAME FOR MEDICAID}, or {STATE CHIP NAME}/{PROGRAM NAME FROM HX160/HX270}, the program sponsored by a state or local government agency which provides hospital and physician benefits,} } {are/is}/{were/was} } {PERSON 1}, {PERSON 2}, {PERSON 3}, {PERSON 4}, {PERSON N} enrolled in an HMO, that is a Health Maintenance Organization {between {START DATE} and {END DATE}}?

[With an HMO, you must generally receive care from HMO physicians. If another doctor is seen, the expense is not covered unless you were referred by the HMO, or there was a medical emergency.]

In subsequent rounds, for persons who had been previously identified as covered by Medicaid, the respondent was asked whether the name of the person's insurance plan had changed since the previous interview. An affirmative response triggered the previous set of questions about managed care (name on list of Medicaid HMOs or signed up with an HMO).

In each round, the variables MCDHMOrr have five possible values:

- 1 The person was covered by a Medicaid/SCHIP HMO.
- 2 The person was covered by Medicaid/SCHIP but the plan was not an HMO.
- 3 The person was not covered by Medicaid/SCHIP.
- 15 The person was covered by Medicaid/SCHIP but the plan type cannot be computed.
- 1 The person was out-of-scope.

Medicaid/SCHIP Gatekeeper Plans

If a person did not belong to a Medicaid/SCHIP HMO, a third question was used to determine whether the person was in a gatekeeper plan. The variables MCDMCrr were set to "Yes" if the respondent provided an affirmative response to the following question:

{Does/Between {START DATE} and {END DATE}, did} {Medicaid},{STATE NAME FOR MEDICAID}, or {STATE CHIP NAME}/{PROGRAM NAME FROM HX160/HX270}, the program sponsored by a state or local government agency which provides hospital and physician benefits,} require {PERSON 1}, {PERSON 2}, {PERSON 3}, {PERSON 4}, {PERSON N} to sign up with a certain primary care doctor, group of doctors, or with a certain clinic which they must go to for all of their routine care?

PROBE: Do not include emergency care or care from a specialist they were referred to.

In each round, the variables MCDMCrr have five possible values:

- 1 The person was covered by a Medicaid/SCHIP gatekeeper plan.
- 2 The person was covered by Medicaid/SCHIP, but it was not a gatekeeper plan.
- 3 The person was not covered by Medicaid/SCHIP.
- 15 The person was covered by Medicaid/SCHIP but the plan type cannot be computed.
- 1 The person was out-of-scope.

Private Managed Care Plans

Persons with private insurance were identified from their responses to questions in the health insurance section of the MEPS questionnaire. In some cases, persons were assigned private insurance as a result of comments collected during the interview, but data editing was minimal. As a consequence, most persons with private insurance were asked about the characteristics of their plan, and their responses were used to identify HMO and gatekeeper plans.

Private HMOs

Persons with private insurance were classified as being covered by an HMO if they met any of the three following conditions:

1. The person reported that their insurance was purchased directly through an HMO,
2. The person reporting private insurance coverage identified the type of insurance company as an HMO, or
3. The person answered “Yes” to the following question:

{Is/Was} {your/{POLICYHOLDER}'s} {NAME OF INSURER} an HMO {as of {END DATE}}? {When answering this question, do not consider {your/his/her} insurance through Medicare.}

[With an HMO, you must generally receive care from HMO physicians. For other doctors, the expense is not covered unless you were referred by the HMO or there was a medical emergency.]

In subsequent rounds, policyholders were asked whether the name of their insurance plan had changed since the previous interview. An affirmative response triggered the detailed question about managed care (i.e., was the insurer an HMO).

Some insured persons have more than one private plan. In these cases, if the policyholder identified any plan as an HMO, the variables PRVHMOrr were set to “Yes.” If a person had multiple plans and one or more were identified as not being an HMO and the other(s) had missing plan type information, the person-level variable was set to missing. Additionally, if a person had multiple plans and none were identified as an HMO, the person-level variable was set to “No.” In each round, the variables PRVHMOrr have five possible values:

- 1 The person was covered by a private HMO.
- 2 The person was covered by private insurance, but it was not an HMO.
- 3 The person was not covered by private insurance.
- 15 The person was covered by private insurance, but the plan type cannot be computed.
- 1 The person was out-of-scope.

Dental and Prescription Drug Private Insurance Variables (DENTIN31-PMDINSyy)

Dental Private Insurance Variables

Round-specific variables (DENTINrr) are provided that indicate the person was covered by a private health insurance plan that included at least some dental coverage for each round of 2021. It should be noted that the information was elicited from a pick-list, code-all-that-apply, question that asked what type of health insurance a person obtained through an establishment. The list included: hospital and physician benefits including coverage through an HMO, Medigap coverage, vision coverage, dental, and prescription drugs. It is possible that some dental coverage provided by hospital and physician plans was not independently enumerated in this question. Users should also note that persons with missing information on dental benefits for all reported private plans and those who reported that they did not have dental coverage for one or more plans but had missing information on other plans are coded as not having private dental coverage. Persons with reported dental coverage from at least one reported private plan were coded as having private dental coverage.

DENTIN53 reflects coverage for all of Panel 26 Round 3, all of Panel 25 Round 5, all of Panel 24 Round 7, and all of Panel 23 Round 9 where the end reference year for Panels 24 and 26 could extend into 2022. DENTIN31 for Panel 25 Round 3, Panel 24 Round 5, and Panel 23 Round 7 reflects coverage in 2020 and 2021 since the reference period for all three rounds spans both years. A second version of these dental coverage indicators was built to reflect only current year coverage (DNTINSrr).

Prescription Drug Private Insurance Variables

Round-specific variables (PMEDINrr) are provided that indicate the person was covered by a private health insurance plan that included at least some prescription drug insurance coverage for each round of 2021. It should be noted that the information was elicited from a pick-list, code-all-that-apply, question that asked what type of health insurance a person obtained through an establishment. The list included: hospital and physician benefits including coverage through an HMO, Medigap coverage, vision coverage, dental, and prescription drugs. It is possible some prescription drug coverage provided by hospital and physician plans was not independently enumerated in this question. Persons with reported prescription drug coverage from at least one reported private plan were coded as having private prescription drug coverage. Users should note that persons with missing information on prescription drug benefits for all reported private plans and those who reported that they did not have prescription drug coverage for one or more plans but had missing information on other plans are coded as not having private prescription drug coverage.

PMEDIN53 reflects coverage for all of Panel 26 Round 3, all of Panel 25 Round 5, all of Panel 24 Round 7, and all of Panel 23 Round 9, where the end reference year for Panels 24 and 26 could extend into 2022. PMEDIN31 for Panel 25 Round 3, Panel 24 Round 5, and Panel 23 Round 7 reflects coverage in 2020 and 2021 since the reference period for all three rounds spans both years. A second version of these prescription drug coverage indicators was built to reflect only current year coverage (PMDINSrr).

Medical Debt Variables (PROBPY42 - PYUNBL42)

Questions relating to medical debt were asked in the health insurance section. Respondents in Round 2, Round 4, Round 6, or Round 8 were asked questions HX770 (“In the past 12 months did anyone in the family have problems paying or were unable to pay any medical bills?”), HX780 (“Does anyone in your family currently have any medical bills that are being paid off over time?”), and HX790 (“Does anyone in your family currently have any medical bills that you are unable to pay at all?”). The corresponding constructed variables PROBPY42, CRFMPY42, and PYUNBL42 are included in this file. PROBPY42 was set to 1 (Yes) if the respondent indicated that someone in their family had problems paying or were unable to pay any medical bills. Additional questions ascertained if anyone in the family currently had medical bills that were being paid off over time (CRFMPY42), and if anyone in the family currently had any medical bills that were unable to be paid at all (PYUNBL42). If the respondent indicated that someone in their family currently has any medical bills that are being paid off over time, then CRFMPY42 was set to 1 (Yes). Note that if the respondent indicates that no one in their family had problems paying medical bills, then PYUNBL42 is set to -1 (Inapplicable).

Prescription Drug Usual Third Party Payer Variables (PMEDUP31-PMEDPY53)

Round-specific variables are provided that indicate whether the sample member had a usual third party payer for prescription medications (PMEDUPrr), and if so, what type of payer (PMEDPYrr). These questions were asked only of sample members who reportedly had at least one prescription medication purchase in the round. In each interview, if the sample member reportedly had a third party payer, then the respondent was asked the name of the sample member’s usual third party payer. These responses were coded into the following source of payment categories in PMEDPYrr: Private Insurance, Medicare, Medicaid, VA/CHAMPVA, TRICARE, State/Local Government, and Other. Users should note that these questions were asked in the Prescribed Medicines (PM) section of the questionnaire, and that no attempt was made to reconcile the responses with information collected in the health insurance section of the questionnaire. In particular, respondents may report the names of private entities (such as insurance companies) that contract with public programs, and these may be coded as private insurance instead of the public programs.

2.5.11 Utilization, Expenditures, and Sources of Payment Variables (TOTCH21-RXOSR21)

The MEPS Household Component (HC) collects data in each round on use and expenditures for office- and hospital-based care, home health care, dental services, vision aids, and prescribed medicines. Data were collected for each sample person at the event level (e.g., doctor visit, hospital stay) and summed across Rounds 7 - 9 for Panel 23 (excluding 2020 events covered in Round 7), Rounds 5 - 7 for Panel 24 (excluding 2020 events covered in Round 5 and 2022 events covered in Round 7), Rounds 3 - 5 for Panel 25 (excluding 2020 events covered in Round 3) and across Rounds 1 - 3 for Panel 26 (excluding 2022 events covered in Round 3) to produce the annual utilization and expenditure data for 2021. In addition, the MEPS Medical Provider Component (MPC) is a follow-back survey that collected data from a sample of medical providers and pharmacies that were used by sample persons in 2021. Expenditure data collected

in the MPC are generally regarded as more accurate than information collected in the HC and were used to improve the overall quality of MEPS expenditure data in this file (see below for description of methodology used to develop expenditure data).

This file contains utilization and expenditure variables for several categories of health care services. In general, there is one utilization variable (based on HC responses only), 11 expenditure variables (derived from both HC and MPC responses), and one charge variable for each category of health care service. The utilization variable is typically a count of the number of medical events reported for the category. The 11 expenditure variables consist of an aggregate total payments variable and 10 main component source of payment category variables (see below for description of source of payment categories). Expenditure variables for all categories of health care combined are also provided. These variables generally represent a full year of use and expenditures. However, for persons who were not in scope for the entire year, these variables reflect only the period of eligibility.

The table in Appendix 3 provides an overview of the utilization and expenditure variables included in this file. For each health service category, the table lists the corresponding utilization variable(s) and provides a general key to the expenditure variable names (11 per service category). The first three characters of the expenditure variable names reflect the service category (except only two characters for prescription medicines) while the subsequent three characters (***) in table reflect the naming convention for the source of payment categories described below (except only two characters for Veterans Administration). The last two positions of all utilization and expenditure variable names reflect the survey year (i.e., 21). More details are provided on the utilization and expenditure variables in Section 2.5.11 below.

Expenditures Definition

Expenditures on this file refer to what is paid for health care services. More specifically, expenditures in MEPS are defined as the sum of direct payments for care provided during the year, including out-of-pocket payments and payments by private insurance, Medicaid, Medicare, and other sources. Payments for over-the-counter drugs are not collected in MEPS. Indirect payments not related to specific medical events, such as Medicaid Disproportionate Share and Medicare Direct Medical Education subsidies, are also not included.

The definition of expenditures used in MEPS is somewhat different from the 1987 NMES and 1977 NMCES surveys where charges rather than sum of payments were used to measure expenditures. This change was adopted because charges became a less appropriate proxy for medical expenditures during the 1990s due to the increasingly common practice of discounting charges. Another change from the two prior surveys is that charges associated with uncollected liability, bad debt, and charitable care (unless provided by a public clinic or hospital) are not counted as expenditures because there are no payments associated with those classifications.

While the concept of expenditures in MEPS has been operationalized as payments for health care services, variables reflecting charges for services received are also provided on the file (see below). Analysts should use caution when working with the charge variables because they do not typically represent actual dollars exchanged for services or the resource costs of those services.

Data Sources on Expenditures

The expenditure data included on this file were derived from the MEPS Household and Medical Provider Components. Only HC data were collected for non-physician visits, dental and vision services, other medical equipment and services, and home health care not provided by an agency. Data on expenditures for care provided by home health agencies were only collected in the MPC. In addition to HC data, MPC data were collected for a sample of office-based visits to physicians (or medical providers supervised by physicians), hospital-based events (e.g., inpatient stays, emergency room visits, and outpatient department visits), and prescribed medicines. For these types of events, MPC data were used if complete; otherwise, HC data were used if complete. Missing data for events where HC data were not complete and MPC data were not collected or complete were derived through an imputation process (see below).

A series of logical edits were applied to both the HC and MPC data to correct for several problems including, but not limited to, outliers, copayments or charges reported as total payments, and reimbursed amounts that were reported as out-of-pocket payments. In addition, edits were implemented to correct for misclassifications between Medicare and Medicaid and between Medicare HMOs and private HMOs as payment sources. Data were not edited to ensure complete consistency between the health insurance and source of payment variables on the file.

Imputation for Missing Expenditures and Data Adjustments

Expenditure data were imputed to 1) replace missing data, 2) provide estimates for care delivered under capitated reimbursement arrangements, and 3) to adjust household-reported insurance payments because respondents were often unaware that their insurer paid a discounted amount to the provider. This section contains a general description of the approaches used for these three situations. A more detailed description of the editing and imputation procedures is provided in the documentation for the MEPS event-level files.

The predictive mean matching (PMM) imputation method was used to impute missing expenditures. This procedure uses regression models (based on events with completely reported expenditure data) to predict total expenses for each event. Then, for each event with missing payment information, a donor event with the closest predicted payment with the same pattern of expected payment sources as the event with missing payment was used to impute the missing payment value.

The general approach that was used to impute missing expenditure data on prescribed medicines is described in Section 2.5.11 below.

Payments under capitated arrangement and, in some cases, public clinics are not tied to individual episodes of medical care. Therefore, expenditures for medical care covered under capitated arrangement or for some services delivered in public settings were imputed. Using a weighted sequential hot-deck procedure, events covered under capitated arrangements were imputed using donor events covered by a managed care organization but paid on a fee-for-service basis. For other events, including public clinics, expenditures were imputed using the PMM method where selected predictor variables were used to predict expenditures and match recipient and donor events.

An adjustment was also applied to some HC-reported expenditure data because an evaluation of matched HC/MPC data showed that respondents who reported that charges and payments were equal were often unaware that insurance payments for the care had been based on a discounted charge. To compensate for this systematic reporting error, a weighted sequential hot-deck imputation procedure was implemented to determine an adjustment factor for HC-reported insurance payments when charges and payments were reported to be equal. As for the other imputations, selected predictor variables were used to form groups of donor and recipient events for the imputation process.

Methodology for Flat Fee Expenditures

Most of the expenditures for medical care reported by MEPS participants are associated with single medical events. However, in some situations one charge covers multiple contacts between a medical provider and patient (e.g., obstetrician services, orthodontia). In these situations (generally called flat or global fees), total payments for the flat or global fee were included if the initial service was provided in 2021. For example, all payments for an orthodontist's fee that covered multiple visits over three years were included if the initial visit occurred in 2021. However, if a visit in 2021 to an orthodontist was part of a flat fee in which the initial visit occurred in 2020, then none of the payments for the flat fee were included.

It is important to note that certain flat fee bundle types reported by healthcare providers (HC) were identified as having a high likelihood of being simple events misidentified as bundle events. To address this, starting in 2021, HC-reported flat fee bundles were considered as flat fees if the bundle consisted only of dental events, or the bundle started in the previous year and also had events in 2021.

Other HC-reported bundles were not allowed as flat fee bundles, and events in these bundles were treated as simple events. HC-reported bundles that included a mix of emergency room and hospitalization events were treated as linked events. All emergency room expenditures were combined with hospital inpatient expenditures. However, provider-reported flat fees were processed in a similar way to prior years.

Zero Expenditures

There are some medical events reported by respondents where the payments were zero. This could occur for several reasons including (1) free care was provided, (2) bad debt was incurred, (3) care was covered under a flat fee arrangement and it was not the initial event of the bundle (see prior section on Methodology for Flat Fee Expenditures), or (4) follow-up visits were provided without a separate charge (e.g., after a surgical procedure). In summary, these types of events have no impact on totals for the person-level expenditure variables contained in this file.

Source of Payment Categories

In addition to total expenditures, variables are provided that itemize expenditures according to the major source of payment categories. These categories are:

1. Out of pocket by patient or patient's family (SLF)--includes any deductible, coinsurance, and copayment amounts not covered by other sources, as well as payments for services and providers not covered by the person's insurance or other sources;

2. Medicare (MCR);
3. Medicaid (MCD);
4. Private Insurance (PRV);
5. Veterans' Administration/CHAMPVA, excluding TRICARE (VA);
6. TRICARE (TRI);
7. Other Federal Sources--includes Indian Health Service, military treatment facilities, and other care provided by the federal government (OFD);
8. Other State and Local Sources--includes community and neighborhood clinics, state and local health departments, and state programs other than Medicaid (STL);
9. Worker's Compensation (WCP);
10. Other Unclassified Sources--includes sources such as automobile, homeowner's, liability, and other miscellaneous or unknown sources (OSR).

Prior to 2019, for cases where reported insurance coverage and sources of payment appear inconsistent, the positive amount from a source inconsistent with reported insurance coverage was moved to one or both of the source categories Other Private and Other Public. Beginning in 2019, this step is removed and the apparent inconsistency between the payment sources and insurance coverage is allowed to remain - the amounts are not moved to Other Private and Other Public categories any more. The two source of payment categories, Other Private and Other Public, are no longer available. Some inconsistencies arise from either misreporting of health insurance coverage or sources of payment. However, apparent inconsistencies may also have logical explanations. For example, private insurance coverage in MEPS is defined as having a major medical plan covering hospital and physician services. If a MEPS sample person did not have such coverage but had a single service type insurance plan (e.g., dental insurance) that paid for a particular episode of care, those payments may be classified as "other private." Some of the "other public" payments may stem from confusion between Medicaid and other state and local programs or may be for persons who were not enrolled in Medicaid, but were presumed eligible by a provider who ultimately received payments from the program.

The naming conventions used for the source of payment expenditure variables are shown in parentheses in the list of categories above and in the key to the attached table in Appendix 3. In addition, total expenditure variables (EXP in key) based on the sum of the 10 source of payment variables above are provided.

Charge Variables

In addition to the expenditure variables described above, a variable reflecting total charges is provided for each type of service category (except prescribed medicines). This variable represents the sum of all fully established charges for care received and usually does not reflect actual payments made for services, which can be substantially lower due to factors such as negotiated discounts, bad debt, and free care (see above). The weighted sequential hot-deck

procedure was used to impute the missing total charges. The naming convention used for the charge variables (TCH) is also included in the key to the attached table in Appendix 3. The total charge variable across services (TOTTCH21) excludes prescribed medicines.

Utilization and Expenditure Variables by Type of Medical Service

The following sections summarize definitional, conceptual, and analytic considerations when using the utilization and expenditure variables in this file. Separate discussions are provided for each MEPS medical service category. There is also a discussion in the section dealing with analyses of trends using MEPS data (Section 3.12).

Medical Provider Visits (i.e., Office-Based Visits)

Medical provider visits consist of encounters that took place primarily in office-based settings and clinics. Care provided in other settings such as a hospital, nursing home, or a person's home are not included in this category.

The total number of office-based visits reported for 2021 (OBTOTV21) as well as the number of such visits to physicians (OBDRV21) are contained in this file.

Expenditure variables associated with all medical provider visits and physician visits can be identified using the attached table in Appendix 3.

Hospital Events

Separate utilization variables for hospital care are provided for each type of setting (outpatient department, emergency room, and inpatient stays) along with three expense variables per setting: one for basic hospital facility expenses, one for payments to physicians who billed separately for services provided at the hospital (referred to as "separately billing doctor" or SBD expenses) and one that aggregates the facility and SBD expenses (aggregated variable not included in files prior to 2007).

Hospital facility expenses include all expenses for direct hospital care, including room and board, diagnostic and laboratory work, x-rays, and similar charges, as well as any physician services included in the hospital charge. SBD expenses typically cover services provided to patients in hospital settings by providers like radiologists, anesthesiologists, and pathologists, whose charges are often not included in hospital bills.

Hospital Outpatient Visits

Variables for the total number of reported visits to hospital outpatient departments in 2021 (OPTOTV21) as well as the number of outpatient department visits to physicians (OPDRV21) are contained in this file.

Expenditure variables (both facility and SBD) associated with all medical provider visits and physician visits can be identified using the attached table in Appendix 3.

Hospital Emergency Room Visits

The variable ERTOT21 represents a count of all emergency room visits reported for the survey year. Expenditure variables associated with ERTOT21 are identified in the attached table in Appendix 3. It should be noted that for emergency room visits that immediately preceded an

inpatient stay, the facility expenditures associated with the emergency room visits are included in the inpatient expenditures. To avoid double counting, these emergency room visits resulted in \$0 facility expenditures (but there still may be associated SBD expenses). However, these \$0 emergency room visits are still counted as separate visits in the utilization variable ERTOT21.

Hospital Inpatient Stays

Two measures of total inpatient utilization are provided on the file:

- IPDIS21 is the total number of hospital discharges.
- IPNGTD21 is the total number of nights associated with these discharges. Please note that the variable IPNGTD21 is an imputed version of the IPNGT21 variable released earlier on HC-228. For the FY 2021 data, the number of imputed length of stay cases is 7.

Expenditure variables associated with hospital inpatient stays are identified in the attached table in Appendix 3. As described in the previous section, payments associated with emergency room visits that immediately preceded an inpatient stay are included with the inpatient expenditures. In addition, payments associated with healthy newborns are included with expenditures for the mother. Specifically, data used to construct the inpatient utilization and expenditure variables for newborns were edited to exclude stays where the newborn left the hospital on the same day as the mother. This edit was applied because discharges for infants without complications after birth were not consistently reported in the survey, and charges for newborns without complications are typically included in the mother's hospital bill. However, if the newborn was discharged at a later date than the mother was discharged, then the discharge was considered a separate stay for the newborn when constructing the utilization and expenditure variables.

Telehealth Events

The Telehealth (TH) module is asked of all events tagged as TH events by the respondent. As part of the TH module, a question is asked about whether the provider or facility is owned or operated by a hospital. Post-collection, the response to this question is used to reclassify all TH events as either OB or OP. The TH module items were designed to align with the existing OB and OP items to easily allow for reclassifying the event type.

Dental Care Visits

The total number of dental care visits variable (DVTOT21) includes those to any person(s) for dental care including general dentists, dental hygienists, dental technicians, dental surgeons, orthodontists, endodontists, and periodontists.

Home Health Care

In contrast to other types of medical events where data were collected on a per visit basis, information on home health care utilization is collected in MEPS on a per month basis. Variables are provided that indicate the total number of days in 2021 where home health care was received from the following: from any type of paid or unpaid caregiver (HHTOTD21), from agencies, hospitals, or nursing homes (HHAGD21), from self-employed persons (HHINDD21), and from unpaid informal caregivers not living with the sample person (HHINFD21). The number of provider days represents the sum across months of the number of days on which home health

care was received, with days summed across all providers seen. For example, if a person received care in one month from one provider on two different days, then the number of provider days would equal two. The number of provider days would also equal two if a person received care from two different providers on the same day. However, if a person received care from one provider two times on the same day, then the provider days would equal one. These variables were assigned missing values if the number of provider days could not be computed for any month in which the specific type of home health care was received.

Separate expenditure variables are provided for agency-sponsored home health care (includes care provided by home health agencies, hospitals, and nursing homes) and care provided by self-employed persons. The attached table in Appendix 3 identifies the home health care utilization and expenditure variables contained in the file.

Other Medical Equipment and Services

The new CAPI instrument collects round-specific Other Medical (OM) expenditures for all OM types (glasses/contact lenses, ambulance, disposable supplies, and long-term equipment). Please note for disposable supplies, the total charge and out-of-pocket expenditures are collected in a range format. The ranges were replaced with mean dollar amounts of respective expenditures reported in each range in prior years.

Prescribed Medicines

There are one total utilization variable (RXTOT21) and 11 expenditure variables included on the 2021 full-year file relating to prescribed medicines. These 11 expenditure variables include an annual total expenditure variable (RXEXP21) and 10 corresponding annual source of payment variables (RXSLF21, RXMCR21, RXMCD21, RXPRV21, RXVA21, RXTRI21, RXOFD21, RXSTL21, RXWCP21, and RXOSR21). The total utilization variable is a count of all fills and refills of prescribed medications obtained during 2021. The total expenditure variable sums all amounts paid out-of-pocket and by third party payers for each prescription obtained in 2021. No variables reflecting charges for prescription medicines are included because a large proportion of respondents to the MEPS pharmacy component survey did not provide charge data (see below).

Prescribed Medicines Data Collected

Data regarding prescription drugs were obtained through the household questionnaire and a pharmacy component survey. During each round of the MEPS HC, all respondents were asked to supply the name of any prescribed medication they or their family members purchased or otherwise obtained during that round. For each medication and in each round, the following information was collected: the number of times the prescription drug was obtained or purchased. In the round each medication was first reported, the following information was collected: the name(s) of any health conditions the medication was prescribed for, and the year and month which the person first used the medication. Respondents were also asked the names, addresses, and types of pharmacies that filled the household's prescriptions. Payment information was collected in the pharmacy component survey.

Pharmacy providers identified by the household were contacted by telephone in the pharmacy component if permission to release their pharmacy records was obtained in writing from the person with the prescription. The signed permission forms were provided to the various establishments prior to making any requests for information. Each establishment was informed

of all persons participating in the survey that had prescriptions filled there in 2021 and a computerized printout containing information about these prescriptions was sought. For each medication listed, the following information was requested: national drug code (NDC), medication name, strength of medicine (amount and unit), quantity (package size and amount dispensed), and payments by source.

Information about diabetic supplies and equipment, such as syringes and test strips, were reported in the prescription drug section of MEPS and use of and expenditures for these items are included in the person-level prescribed medicine variables.

Prescribed Medicines Data Editing and Imputation

The general approach to preparing the household prescription data for this file was to utilize the pharmacy component prescription data to assign expenditure values to the household drug mentions. For those with Pharmacy Component data, a matching program was adopted to link pharmacy component drugs and the corresponding drug information to household drug mentions. To improve the quality of these matches, all drugs on the household and pharmacy files were coded based on the medication names provided by the household and pharmacy, and when available, the national drug code (NDC) provided in the pharmacy survey. Considerable editing was done prior to the matching to correct data inconsistencies in both data sets, fill in missing data, and correct outliers on the pharmacy file.

Drug price per unit outliers were analyzed on the pharmacy file by first identifying the national average drug acquisition cost (NADAC) per unit, wholesale acquisition unit cost (WAUC), and average wholesale unit price (AWUP) of the drug by linkage through the NDC to secondary databases. In general, prescription drug unit prices were deemed to be outliers by comparing unit prices reported in the pharmacy database to the NADAC per unit and were edited, as necessary. Prior to 2020, AWUP was the benchmark used to identify outlier prices for prescription medications in the PC. Beginning with the 2007 data, the rules used to identify outlier prices relative to the AWUP changed. New outlier thresholds were established based on the distribution of the ratio of retail unit prices relative to the AWUP in the 2006 MarketScan Outpatient Pharmaceutical Claims database.

Starting with the 2008 Prescribed Medicine file, improvements in the data editing changed the distribution of payments by source: (1) more spending on Medicare beneficiaries is by private insurance, rather than Medicare, and (2) there are less out-of-pocket payments and more Medicaid payments among Medicaid enrollees. Beginning with the 2009 data, another change affected the data for Medicare beneficiaries with both Part D and Medicaid coverage: reported Medicaid and other state and local program payments were no longer edited to be Medicare payments. Beginning with the 2017 data, changes in the price imputation procedures for specialty drugs with missing payment information resulted in higher total prescribed medicines expenditures.

Beginning with the 2020 data, the rules used to identify outlier prices for prescription medications in the PC were improved based on newer price benchmarks and analyses (Ding & Hill, 2022). New outlier thresholds were established based on the distribution of the ratio of retail unit prices relative to the NADAC per unit, collected for the Centers for Medicare and Medicaid Services. When the NADAC per unit is not available, then the WAUC is used, and if

neither are available, the AWUP is used. AWUP and WAUC are list prices, not averages, so the NADAC per unit better reflects the prices paid for drugs, and as a result the prices paid for generics are lower in the 2020 data, compared with the 2019 data, and fewer generic fills have third party payments.

For Round 3, Panel 24 Round 5, and Round 7, which span two years, drug mentions in those rounds were allocated between the years based on the following information: the number of times the respondent said the drug was purchased in the respective year, the year the person started taking the drug, the length of the person’s round, the dates of the person’s round, and the number of fills of that drug for that person in the round.

Collapsed Source of Payment Variables

Two additional source of payment variables are included for each health care service category as a convenience to data users since they are common analytic groupings of the payment sources. The first (**PTR21 series) is the sum of the private and Tricare payer categories (i.e., **PTR21=**PRV21+**TRI21). The second (**OTH21 series) is the sum of the least common source of payment categories including: 1) other federal (**OFD21), 2) state and local (**STL21), and 3) other sources (**OSR21). Since the **PTR21 and **OTH21 variable series represent combined totals of existing individual source of payment variables, analysts should exercise caution to avoid inappropriate double counting of expenditures when working with these variables.

2.5.12 Changes in Variable List

Variables were added and removed from the file due to changes in the questions asked in 2021 relative to prior years. The MEPS HC questionnaires can be found on the MEPS website.

Following is a list of changes to the variable list for the 2021 full-year consolidated data file.

Added

- BOOSTERSHOT53
- CVDLAYCA31
- CVDLAYCA42
- CVDLAYDN31
- CVDLAYDN42
- CVDLAYPM31
- CVDLAYPM42
- CVVACCINE42
- CVVACCINE53
- SDAFRDHOME
- SDCHURCH
- SDCLUBORG
- SDCMPM
- SDCMPY
- SDCOMM
- SDCOMPAN
- SDDAYEXER
- SDDEBT
- SDDSCRMDR
- SDDSCRMHS
- SDDSCRMJOB
- SDDSCRMPOL
- SDDSCRMPUB
- SDDSCRMSTR

- SDDSCRMWRK
- SDENICPROD
- SDFAMILY
- SDFRCSXCH
- SDFRIENDS
- SDGETTGT
- SDHLTHFOOD
- SDHMALC
- SDHMBEAT
- SDHMDEPR
- SDHMDIV
- SDHMDRG
- SDHMJAIL
- SDHOME
- SDHURTCHLD
- SDINSCHLD
- SDINSULT
- SDISOL
- SDLATERENT
- SDLATEUTIL
- SDLEFTOUT
- SDLIFE
- SDMEDCARE
- SDMINSEXER
- SDMISSCCLN
- SDNOFOOD
- SDNOTRANS
- SDOHELIG
- SDOHWT21F
- SDPARKS
- SDPAYBASICS
- SDPHYSHURT
- SDPROBCOOK
- SDPROBHEAT
- SDPROBLEAD
- SDPROBLEAKS
- SDPROBMOLD
- SDPROBNONE
- SDPROBPEST
- SDPROBSMKDET
- SDPROX
- SDPUBTRANS
- SDSCREAM
- SDSFCRIME
- SDSHUTUTIL
- SDSTRESS
- SDTCHADLT
- SDTCHCHLD
- SDTHRHARM
- SDTLKPHN
- SDUNEXPEXP
- SDWRRYFD
- TRI21X
- TRI31X
- TRI42X

Added (included in alternating years only, will not be included in 2022):

- ADAPPT42
- ADEXPL42
- ADEZUN42
- ADFFRM42
- ADFHLP42
- ADHECR42
- ADILCR42
- ADILWW42
- ADINSA42
- ADINSB42
- ADINST42
- ADLIST42

- ADNSMK42
- ADOVER42
- ADPRTM42
- ADRESP42
- ADRISK42
- ADRTCR42
- ADRTWW42
- ADSMOK42
- ADSNSP42
- ADSPCL42
- ADTLHW42
- ADUPRO42
- CHAPPT42
- CHEXPL42
- CHEYRE42_M18
- CHHECR42
- CHILCR42
- CHILWW42
- CHLIST42
- CHPRTM42
- CHRESP42
- CHRTCR42
- CHRTWW42
- CHSPEC42_M18
- DADPRO42
- GETTRB42
- HAVFUN42
- HOMEBH42
- KIDPRO42
- MOMPRO42
- NERVAF42
- SCHLBH42
- SCHPRO42
- SIBPRO42
- SPRPRO42
- UNHAP42

Deleted

- OUTFLAG31
- OUTFLAG42
- OUTFLAG53
- TRIEX21X
- TRIEX31X
- TRIEX42X
- TRILI21X
- TRILI31X
- TRILI42X
- TRIPR21X
- TRIPR31X
- TRIPR42X
- TRIST21X
- TRIST31X
- TRIST42X

Deleted (included in alternating years only, will be included in 2022):

- ADAGE42
- ADAPTRT42
- ADBLDS42
- ADBMI42
- ADBNDN42
- ADBPCK42
- ADBRST42
- ADBRTC42
- ADCHLC42
- ADCLNS42
- ADCOLN42
- ADDSCU42

- ADFLST42
- ADGRPTRT42
- ADKALC42
- ADMDVT42
- ADMMGR42
- ADMNTRT42
- ADMOOD42
- ADNNOAP42
- ADNUMDRK42
- ADOFTALC42
- ADOFTB42
- ADONLTRT42
- ADOSTP42
- ADPAP42
- ADPAPG42
- ADPHONTRT42
- ADPNEU42
- ADPROBTRT42
- ADPROS42
- ADPSAG42
- ADQTHP42
- ADQTM42
- ADQTTB42
- ADRATETRT42
- ADRELTRT42
- ADRNK442
- ADRNK542
- ADSCHTRT42
- ADSEX42
- ADSGMD42
- ADSHNG42
- ADSLEEP42
- ADSTAL42
- ADTBAC42
- ADTRTEXP42
- ADTRTHLP42
- ADTRTPD42
- ADUNABTRT42
- ADUTRM42
- ADWGHD42
- ADWTAD42
- BOOST42
- CHBMIX42
- EATHLT42
- HELMET42
- LAPBLT42
- MESHGT42
- MESVIS42
- MESWGT42
- NOSMOK42
- PHYSCL42
- SAFEST42
- TIMALN42
- WHNBST42
- WHNEAT42
- WHNHEL42
- WHNHGT42
- WHNLAP42
- WHNPHY42
- WHNSAF42
- WHNSMK42
- WHNWGT42

2.6 Linking to Other Files

2.6.1 Event and Condition Files

Records on this file can be linked to 2021 MEPS HC public use event and conditions files by the sample person identifier (DUPERSID). The Panel 25 cases on this file (PANEL=25), the Panel 24 cases on this file (PANEL=24), and the Panel 23 cases on this file (PANEL=23) can also be linked back to the 2020 MEPS HC public use event and condition files. In addition, the Panel 23 and Panel 24 cases can be linked back to the 2019 MEPS HC public use event and condition files, and the Panel 23 cases can be linked back to the 2018 MEPS HC public use event and condition files.

2.6.2 National Health Interview Survey

The set of households selected for MEPS is a subsample of those participating in the National Health Interview Survey (NHIS); thus, each MEPS panel can also be linked back to the previous year's NHIS public use data files. For information on obtaining MEPS/NHIS link files please see the [AHRQ website](#).

2.6.3 Longitudinal Analysis

Panel-specific longitudinal files are available for downloading in the data section of the MEPS website. For all four panels (Panel 23, Panel 24, Panel 25, and Panel 26), the longitudinal file comprises MEPS survey data obtained all rounds of the panel and can be used to analyze changes over the entire length of the panel. For Panel 24, a file representing a three-year period will also be established and updated to cover four years with the release of 2022 data. For Panel 23, a file representing a four-year period will be established. Variables in the file pertaining to survey administration, demographics, employment, health status, disability days, quality of care, patient satisfaction, health insurance, and medical care use and expenditures were obtained from the MEPS full-year Consolidated files from the years covered by each panel.

For more details or to download the data files, please see Longitudinal Weight Files at the [AHRQ website](#).

3.0 Survey Sample Information

3.1 Discussion of Pandemic Effects on Quality of 2021 MEPS Data

3.1.1 Summary

The challenges associated with MEPS data collection in 2020 after the onset of the COVID-19 pandemic continued into 2021. The major modifications to the standard MEPS study design remained in effect, permitting data to be collected safely but with accompanying concerns related to the quality of the data obtained. These data quality issues are discussed below. The suggestion made in the documentation for the [FY2020 MEPS Consolidated Data File](#) (as well as for most federal major in-person surveys conducted in 2021 and 2020) still holds. Researchers are counseled to take care in the interpretation of estimates based on data collected from these two calendar years. This includes the comparison of such estimates to those of other years and corresponding trend analyses.

3.1.2 Overview

Section 3.1 of the documentation for the [2020 Full Year Consolidated Data File](#) provides a general discussion of the impact of the COVID-19 pandemic on several other major in-person federal surveys as well as on MEPS. In addition, it offers a detailed look at how MEPS was modified to permit safe data collection and the development of useful estimates at a time when the way the U.S. health care system functioned underwent many transformations to meet population needs.

In this corresponding 2021 document, focus is placed mostly on MEPS data quality in 2021. However, it also includes how data quality issues related to the two federal surveys most closely connected to it, the National Health Interview Survey (NHIS) carried out by the National Center for Health Statistics (NCHS) and the Current Population Survey (CPS) carried out by the Census Bureau, have an impact on the data quality issues of MEPS.

Specifically, the following discussion describes: 1) data quality issues experienced by the NHIS and CPS that affect MEPS; 2) modifications to the MEPS sample design in 2021 due to the continuing pandemic; and 3) potential data quality issues in the FY 2021 MEPS data related to the COVID-19 pandemic.

3.1.3 Data Quality Issues for MEPS in 2021 Directly Associated with Data Quality Concerns for the NHIS and CPS

Households fielded for Round 1 of MEPS in each year have been selected as a subsample from among the NHIS responding households from the prior year. The MEPS first year panel in 2021 was Panel 26. The households fielded for MEPS in Round 1 of Panel 26 were thus selected from NHIS responding households in 2020. It is important to note here that the NHIS households eligible for use in MEPS are restricted to the first three quarters of the NHIS as the fourth quarter

households cannot be made available in time for MEPS data collection early in the next calendar year.

The onset of the pandemic in 2020 at a national level took place in mid-March of that year, when the NHIS data collection for the first quarter of 2020 was virtually completed and that of the second quarter was about to begin. The NHIS had to make a rapid transition from in-person to telephone interviewing, to attempt to gather NHIS data for the second quarter of 2020. While NCHS was able to make the transition, assessments made by NCHS at the time indicated a much lower response rate than is typically experienced during Quarter 2 and the quality of Quarter 2 data was of particular concern. NCHS thus modified the 2020 NHIS sample design for Quarters 3 and 4. A randomly selected subsample of the sampled housing units originally selected for fielding in Quarters 3 and 4 of 2020 was removed from the sample to be fielded. This reduced sample for Quarters 3 and 4 was then enhanced by randomly selecting responding households from the 2019 NHIS for interviewing in 2020 as well. In consideration of the data quality issues and sample design modifications associated with the 2020 NHIS, the MEPS sample design for FY 2021 was modified, as discussed below.

The quality of CPS data is always of particular importance to MEPS as March CPS-ASEC estimates serve as the basis of control totals for the raking component of the MEPS weighting process. These control totals incorporate the following demographic variables: age, sex, race/ethnicity, region, MSA status, educational attainment, and poverty status. The CPS estimates of educational attainment and poverty status used in the development of the FY 2021 MEPS PUFs were of particular concern. Evaluations of these estimates undertaken by the Census Bureau have shown that they suffered from bias due to survey nonresponse, with CPS income estimates being on the high side and the estimate of those under the poverty threshold being on the low side. The impact of these CPS estimates on the quality of MEPS estimates has been carefully considered. The approach used for the MEPS FY 2021 Consolidated PUF sample weights is discussed in Section 3.4

References (Bramlett et al., 2021; Dahlhamer et al., 2021; Lau et al, 2021; Rothbaum & Bee, 2021, 2022; Zuvekas & Kashihara, 2021) discussing the fielding of these surveys during the pandemic and possible bias concerns, can be found in the References section of this document.

3.1.4 Modifications to the MEPS HC 2021 Sample Design

The MEPS sample is usually selected from Quarters 1 to 3 of NHIS Panels 1 and 3. However, 2020 NHIS Quarter 2 data quality and sample size issues needed to be considered in planning for selecting and fielding a MEPS sample for 2021. It became clear that it would be prudent to modify the 2021 MEPS sample design for MEPS Panel 26 by replacing 2020 NHIS Quarter 2 with another quarter from 2020 NHIS Panels 2 and 4. Action had to be taken immediately because the MEPS sample selection from NHIS responding households begins in the late summer/early fall of each year.

AHRQ contacted NCHS, reviewing the various issues and asking if responding households in NHIS Panels 2 and 4 from Quarter 1 of 2020 could be made available for MEPS sample selection. Virtually all of these households were interviewed in-person prior to the major onset of

the pandemic, so the Quarter 1 response rates for all four NHIS panels were consistent with prior years and the data quality issues associated with the pandemic could be avoided. NCHS was fully supportive of this approach and made NHIS Panels 2 and 4 for Quarter 1 available for use by MEPS. Thus, for MEPS Panel 26, the NHIS responding households subsampled for MEPS were selected among all NHIS responding households in all four NHIS panels in Quarter 1 as well as those responding in Quarter 3 of NHIS Panels 1 and 3.

As an adjunct to this modification, it was decided to take advantage of the additional PSUs (sampled localities) available from NHIS Panels 2 and 4 and appearing in the MEPS sample for the first time. State level estimation is of interest to MEPS, and the added PSUs would serve to increase the precision of variance estimation for state level estimates. State estimates that would be expected to benefit the most from these added PSUs were the “middle-sized” states where sample sizes were marginal but usually not enough PSUs for variance estimation. The largest states already had large sample sizes while precision for the smallest states would remain low. As a result, from Quarter 1 of 2020 NHIS Panels 2 and 4, the MEPS sample focused on oversampling the “middle-sized” states rather than Asians, Blacks, and Hispanics.

Finally, it was decided to collect 2021 data for nine rounds of Panel 23 and seven rounds of Panel 24, so that these two panels will contribute to MEPS estimates for four calendar years for Panel 23 and three calendar years for Panel 24. In so doing, the number of respondents to MEPS will be kept at a relatively high level despite the decline in response rates due to the pandemic. The MEPS FY 2021 PUF records thus consist of data obtained from the following MEPS Panels and corresponding rounds: Panel 23, Rounds 7-9; Panel 24, Rounds 5-7; Panel 25, Rounds 3-5; and Panel 26, Rounds 1-3.

3.1.5 Data Quality Issues for MEPS for FY 2021

Three sources of potential bias were identified for MEPS for FY 2020: long recall period for Round 6 of Panel 23; switching from in-person to telephone interviewing which likely had a larger impact on Panel 25; and the impact of CPS bias on the MEPS weights. A number of statistically significant differences were found between panels for FY 2020. Those findings are discussed in the documentation for the 2020 Full Year Consolidated Data File, MEPS HC-224.

With this in mind, there were a number of uncertainties for FY 2021 warranting examination. Would Panel 23 data quality increase substantially once the issue of an extensive recall period was eliminated? Would the switching from in-person to telephone interview in Round 1 continue to impact Panel 25 estimates? Since Panel 26 was the first year MEPS panel in 2021, would Panel 26 estimates tend to be different than those of the other three panels?

Analyses undertaken to examine the quality of MEPS FY 2021 data appearing on the Consolidated PUF have been focused on the comparison of health care utilization and health insurance status distribution (some private insurance, some public insurance, no health insurance) for the MEPS target population between the panels fielded. These comparisons were undertaken for the full sample and the three age groups of 0-17, 18-64, and 65+.

Comparisons of health care utilization data for Panel 24 and Panel 23 found no statistically significant differences in any of the analysis variables, suggesting that Panel 23 data quality

increased substantially once the issue of an extensive recall period was eliminated. Comparisons of Panel 26 and Panel 24 health care utilization data did not detect any notable differences among event types of a salient nature.

Similar to FY 2020, comparisons of Panel 25 health care utilization data with those from other panels found that the difference in estimates reached statistical significance for several event types, with those from Panel 25 generally being higher. However, the gap between the Panel 25 estimates and other panels has shown a sizeable decrease.

Another set of analyses compared estimates based on the four panels pooled together using the FY 2021 Consolidated PUF weight compared to the estimates based on the individual panel data and weights. Analyses indicated that event estimates combining data from all four panels show little difference from event estimates of the individual panels.

The analyses undertaken for health insurance status distribution suggest no major differences between the four panels for various event type estimates. Even though slight differences were observed with Panel 25 (e.g., the distribution associated with the age range 18-64 showed a higher percentage of all public insurance compared to the other three panels while those at least 65 years of age showed a lower percentage of some private insurance compared to the other three panels), no statistically significant differences were detected.

Ongoing work is also comparing trends in MEPS utilization estimates to external data sources including preliminary 2021 data from the Healthcare Cost and Utilization Project, Medicare fee for service claims, Medicare Advantage encounter data, Medicaid T-MSIS data, and private insurance claims. Preliminary analyses suggest that inpatient hospital stays and emergency room utilization in 2021 were lower than expected when comparing trends in the MEPS to trends in these other sources. However, exact comparisons are difficult to make because of differences in populations covered, definitional differences, and other differences between MEPS and administrative data sources. Nevertheless, researchers are cautioned to take particular care in the interpretation of estimates based on inpatient hospital stays and emergency room visits. No such concerns were noted with other types of health care events in the comparisons to external data sources. AHRQ continues to investigate the reporting of all health care events in the MEPS throughout the pandemic.

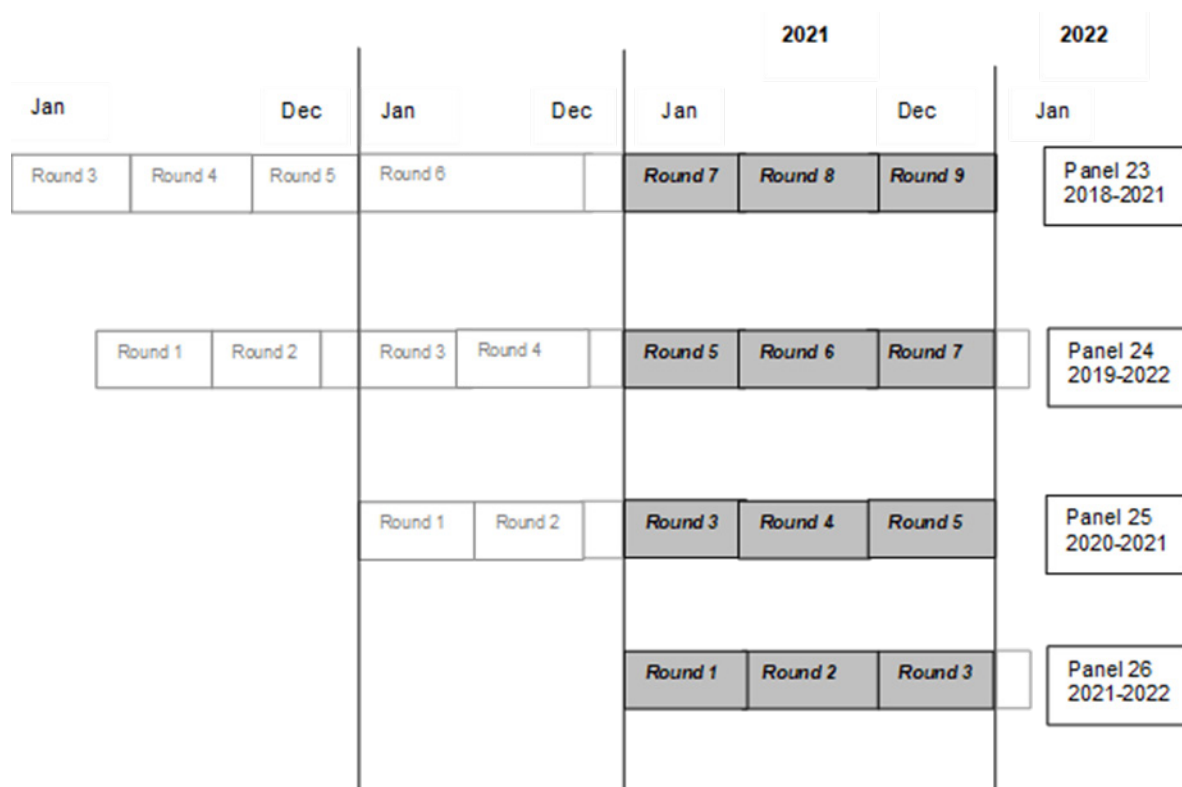
3.1.6 Discussion and Guidance

The various actions taken in the development of the person-level weights for the MEPS FY 2021 Consolidated PUF were designed to limit the potential for bias in the data due to changes in data collection and response bias. However, evaluations of MEPS data quality in 2021, consistent with those of other Federal surveys fielded in 2021, suggest that users of the MEPS FY 2021 Consolidated PUF should exercise caution when interpreting estimates and assessing analyses based on these data as well as in comparing 2021 estimates to those of prior years.

3.2 Background on Sample Design and Response Rates

The MEPS is designed to produce estimates at the national and regional level over time for the civilian, noninstitutionalized population of the United States and some subpopulations of interest. The data in this public use file pertain to calendar year 2021. The data were collected in Rounds 1, 2, and 3 for MEPS Panel 26, Rounds 3, 4, and 5 for MEPS Panel 25, Rounds 5, 6, and 7 for MEPS Panel 24, and Rounds 7, 8, and 9 for MEPS Panel 23.

As usual, Round 3 for a MEPS panel (this time for Panel 26) has been designed to overlap two calendar years, as illustrated below. However it may be noted that, with the fielding of a third and fourth panel in 2021 (as indicated in the data quality discussion in Section 3.1), the structure of other rounds has changed. Round 7 of Panel 24 and Round 5 of Panel 25 serve the same purpose. Thus, Round 7 was fielded in 2020 and designed to collect data for the remainder of 2020 as well as the period of time from January 1, 2021 up through the date of the Round 7 interview. Round 5 of Panel 24 was designed for the same purpose, collecting data associated with both 2020 and 2021. This was done to permit all three of these panels to provide data for the FY 2021 MEPS data sets as well as those for FY 2020.



Variables convey the same information for this full year consolidated file that has been provided for the full year consolidated files associated with years 1996 - 2020 of MEPS.

A sample design feature shared by Panel 23, Panel 24, and Panel 25 involved the partitioning of the sample domain “Other” (serving as the catchall stratum and consisting mainly of households with “White” members) into two sample domains. This was done for the first time in Panel 16. The two domains distinguished between those households characterized as “complete”

respondents to the NHIS and those characterized as “partial completes.” NHIS “partial completes” typically have a lower response rate to MEPS and for all three MEPS panels the “partial” domain was sampled at a lower rate than the “complete” domain. This approach has served to reduce survey costs, since the “partials” tend to have higher costs in gaining survey participation, but has also increased sample variability due to the resulting increased variability in sampling rates. Starting with Panel 25, the “Other, Partial” domain includes the NHIS households that have provided only a roster of household members. For detailed information on the MEPS sample design, see Chowdhury et al. (2019). This feature was not of particular emphasis in Panel 26, for reasons discussed in Section 3.1.4.

3.2.1 MEPS-Linked to the National Health Interview Survey (NHIS)

Each responding household found in this 2021 MEPS dataset is associated with one of the four separate and overlapping MEPS panels, Panel 23, Panel 24, Panel 25, and Panel 26. These panels consist of subsamples of households participating in the 2017, 2018, 2019, and 2020 NHIS, respectively. The Full Year 2018 PUF was the first one where both MEPS panels reflect the new NHIS sample design first implemented in 2016.

Whenever there is a change in sample or study design, it is good survey practice to assess whether such a change could affect the sample estimates. For example, increased coverage of the target populations with an updated sample design based on data from the latest Census can improve the accuracy of the sample estimates. MEPS estimates have been and will continue to be evaluated to determine if an important change in the survey estimates might be associated with a change in design. Discussion on the potential effects of such MEPS design changes in 2021 appears in the data quality section, 3.1. Background on the two NHIS sample designs of interest here is provided next.

Background on the NHIS Sample Redesign Implemented in 2016

Beginning in 2016, NCHS implemented another new sample design for the NHIS, which differed substantially from the prior design. Each of the 50 states as well as the District of Columbia served as explicit strata for sample selection purposes with the intent of providing the capability of state-level NHIS estimates obtained through pooling across years if the sample size for a single year would result in unreliable estimates. In contrast to the previous design, households in areas with relatively high concentrations of minorities are not oversampled. PSUs are still formed at the county level. However, within sampled PSUs, the clusters of addresses that have been sampled for each year of the NHIS are not in the form of segments (consisting of one or more census blocks) as was done for the previous NHIS designs. For the 2016 NHIS, each such cluster consisted of roughly 25 subclusters selected using random systematic sampling across the full geography of the PSU. Each subcluster is made up of, generally, four nearby addresses, or roughly 100 addresses in all. The number of subclusters per cluster can vary from year to year.

Another major change is that the list of DUs (addresses) was obtained from the Computerized Delivery Sequence File (CDSF) of the U.S. Postal Service, which is a different approach than the standard listing process for area probability samples used in the pre-2016 designs. While

addresses in the CDSF provide very high coverage of most areas of the country, coverage in rural areas can be somewhat lower. For rural areas where this was a concern, address lists were created through the conventional listing process. A [description of the NHIS sample design](#) is provided by NCHS on the NHIS website.

Panel 23 Household Sample Size

A subsample of 9,700 households (occupied DUs) was randomly selected for MEPS Panel 23 from NHIS responding households in 2017, of which 9,694 were fielded for MEPS after the elimination of any units characterized as ineligible for fielding.

Panel 24 Household Sample Size

A subsample of 9,700 households was randomly selected for MEPS Panel 24 from the households responding to the 2018 NHIS, of which 9,684 were fielded for MEPS after the elimination of any units characterized as ineligible for fielding.

Panel 25 Household Sample Size

A subsample of 9,900 households was randomly selected for MEPS Panel 25 from the households responding to the 2019 NHIS, of which 9,888 were fielded for MEPS after the elimination of any units characterized as ineligible for fielding.

Panel 26 Household Sample Size

A subsample of 9,510 households was randomly selected for MEPS Panel 26 from the households responding to the 2020 NHIS, of which all 9,510 were fielded for MEPS after the elimination of any units characterized as ineligible for fielding.

Implications of the New Design on MEPS Estimates

Under the new design, MEPS sampled households reflect the clustering of the NHIS, as described above but to a somewhat lesser degree due to the sampling from NHIS respondents. Due to the spreading of the NHIS sample in small subclusters across the PSU and the sampling limited to only NHIS respondents, the impact of clustering on the variance of MEPS estimates may be more limited. Also, in contrast to the previous design, the NHIS sampling rates at the address level currently do not vary due to oversampling of minorities (although this could change in subsequent years). On balance, the overall variation in sampling rates/weights at the national level for the NHIS is expected to be lower with a corresponding positive impact on the precision of MEPS estimates. However, with a reduction in the sample sizes of minority households, precision levels of MEPS estimates for Asians, Blacks, and Hispanics may be reduced to some extent.

3.2.2 Sample Weights and Variance Estimation

In the dataset “MEPS HC-233: 2021 Full Year Consolidated Data File,” weight variables are provided for generating MEPS estimates of totals, means, percentages, and rates for persons and families in the civilian noninstitutionalized population. The person-level weight variable PERWT21F provided in this file supersedes the corresponding person-level weight variable provided in the 2021 Full Year Population Characteristic File (HC-228). Procedures and considerations associated with the construction and interpretation of person and family-level estimates using these and other variables are discussed below. It should be noted that NCHS has made a modification to the NHIS sample design that has affected the MEPS variance structure. This is discussed in more detail in Section 3.11.1.

3.3 The MEPS Sampling Process and Response Rates: An Overview

For most MEPS panels, a sample representing about three-eighths of the NHIS responding households is made available for use in MEPS. This was the case for MEPS Panel 23, Panel 24, and Panel 25. For Panel 26, the sample made available for use in MEPS represented about five-eighths of the NHIS responding households for the reasons discussed in Section 3.1.4.

Because the MEPS subsampling has to be done soon after NHIS responding households are identified, a small percentage of the NHIS households initially characterized as NHIS respondents are later classified as nonrespondents for the purposes of NHIS data analysis. This actually serves to increase the overall MEPS response rate slightly since the percentage of NHIS households designated for use in MEPS (all those characterized initially as respondents from the NHIS panels and quarters used by MEPS for a given year) is slightly larger than the final NHIS household-level response rate and some NHIS nonresponding households do participate in MEPS. However, as a result, these NHIS nonrespondents who are MEPS participants have no NHIS data available to link with MEPS data. Once the MEPS sample is selected from among the NHIS households, characterized as NHIS respondents, RUs consisting entirely of military personnel are deleted from the sample. Military personnel not living in the same RU as civilians are ineligible for MEPS. After these exclusions, all RUs associated with households, selected from among those identified as NHIS responding households are then fielded in the first round of MEPS.

Table 3.1 shows in Rows A, B, and C the three informational components just discussed. Row A indicates the percentage of NHIS households eligible for MEPS. Row B indicates the number of NHIS households sampled for MEPS. Row C indicates the number of sampled households actually fielded for MEPS (after dropping the military members discussed above and a small number of NHIS households sampled in error). Note that all response rates discussed here are unweighted.

Table 3.1. Sample Size and Unweighted Response Rates for 2021 Full Year File (Panel 26 Rounds 1-3/Panel 25 Rounds 3-5/Panel 24 Rounds 5-7/Panel 23 Rounds 7-9)

Components	Panel 23	Panel 24	Panel 25	Panel 26	2021 Combined
A. Percentage of NHIS households designated for use in MEPS (those initially characterized as responding) *	67.1%	64.3%	65.7%	60.6%	-
B. Number of households sampled from the NHIS	9,700	9,700	9,900	9,510	-
C. Number of Households sampled from the NHIS and fielded for MEPS	9,694	9,684	9,888	9,510	-
D. Round 1 - Number of RUs eligible for interviewing	10,089	10,090	10,152	9,795	-
E. Round 1 - Number of RUs with completed interviews	7,351	7,186	6,265	5,882	-
F. Round 2 - Number of RUs eligible for interviewing	7,492	7,323	5,958	6,045	-
G. Round 2 - Number of RUs with completed interviews	6,960	6,777	4,677	4,799	-
H. Round 3 - Number of RUs eligible for interviewing	7,082	6,890	5,189	4,876	-
I. Round 3 - Number of RUs with completed interviews	6,703	6,289	4,230	4,103	-
J. Round 4 - Number of RUs eligible for interviewing	6,776	6,371-	4,307	-	-
K. Round 4 - Number of RUs with completed interviews	6,522	5,446-	3,685	-	-
L. Round 5 - Number of RUs eligible for interviewing	6,536	5,495-	3,706	-	-
M. Round 5 - Number of RUs with completed interviews	6,383	4,770-	3,278	-	-
N. Round 6 - Number of RUs eligible for interviewing	6,480	4,808	-	-	-
O. Round 6 - Number of RUs with completed interviews	5,120	3,959	-	-	-
P. Round 7 - Number of RUs eligible for interviewing	5,170	4,002	-	-	-

Components	Panel 23	Panel 24	Panel 25	Panel 26	2021 Combined
Q. Round 7 - Number of RUs with completed interviews	4,513	3,500	-	-	-
R. Round 8 - Number of RUs eligible for interviewing	4,548	-	-	-	-
S. Round 8 - Number of RUs with completed interviews	3,984	-	-	-	-
T. Round 9 - Number of RUs eligible for interviewing	3,996	-	-	-	-
U. Round 9 - Number of RUs with completed interviews	3,603	-	-	-	-
Overall annual unweighted response rates P26: A x (E/D) x (G/F) x (I/H) P25: A x (E/D) x (G/F) x (I/H) x (K/J) x (M/L) P24: A x (E/D) x (G/F) x (I/H) x (K/J) x (M/L) x (O/N) x (Q/P) P23: A x (E/D) x (G/F) x (I/H) x (K/J) x (M/L) x (O/N) x (Q/P) x (S/R) x (U/T) Combined: $0.22 \times P23 + 0.22 \times P24 + 0.25 \times P25 + 0.31 \times P26$	22.0% (Panel 23 through Round 9)	20.7% (Panel 24 through Round 7)	19.6% (Panel 25 through Round 5)	24.3% (Panel 26 through Round 3)	21.8%

*Among the panels and quarters of the NHIS allocated to MEPS, the percentage of households that were considered to be NHIS respondents at the time the MEPS sample was selected.

3.3.1 Response Rates

In order to produce annual health care estimates for calendar year 2021 based on the full MEPS sample data from the MEPS Panel 23, Panel 24, Panel 25, and Panel 26, the four panels are combined. More specifically, full calendar year 2021 data collected in Rounds 7 through 9 for the MEPS Panel 23, Rounds 5, 6, and 7 for the MEPS Panel 24, and Rounds 3 through 5 for the MEPS Panel 25 sample are pooled with data from the first three rounds of data collection for the MEPS Panel 26 sample (the general approach is described below).

As mentioned above, all response rates discussed here are unweighted. To understand the calculation of MEPS response rates, some features related to MEPS data collection should be noted. When an RU is visited for a round of data collection, changes in RU membership are identified. Such changes include the formation of student RUs as well as other new RUs created when RU members from a previous round have moved to another location in the U.S. Thus, the number of RUs eligible for MEPS interviewing in a given round is determined after data collection is fully completed. The ratio of the number of RUs completing the MEPS interview in a given round to the number of RUs characterized as eligible to complete the interview for that

round represents the “conditional” response rate for that round expressed as a proportion. It is “conditional” in that it pertains to the set of RUs characterized as eligible for MEPS for that round and thus is “conditioned” on prior participation rather than representing the overall response rate through that round. For example, in Table 3.1, for Panel 26 Round 2 the ratio of 4,799 (Row G) to 6,045 (Row F) multiplied by 100 represents the response rate for the round (79.4 percent when computed), conditioned on the set of RUs characterized as eligible for MEPS for that round. Taking the product of the percentage of the NHIS sample eligible for MEPS (Row A) with the product of the ratios for a consecutive set of MEPS rounds beginning with Round 1 produces the overall response rate through the last MEPS round specified.

The overall unweighted response rate for 2021 for the combined sample after pooling the respondents across the four panels was obtained by computing the product of the compositing factor associated with each panel (discussed in Section 3.4.6 describing the development of the final weight for the FY 2021 Consolidated PUF) and the corresponding overall panel response rate and then summing the four products. Panel 26 represents about 31.5 percent of the combined sample size, Panel 25 represents about 24.7 percent of the combined sample size, and Panel 24 represents about 22.1 percent, and Panel 23 represents the remaining 21.7 percent. Thus, the combined response rate of 21.8 percent was computed as 0.22 times 22.0 (22.0 is the overall Panel 23 response rate through Round 9) plus 0.22 times 20.7 (20.7 is the overall Panel 24 response rate through Round 7) plus 0.25 times 19.6 (19.6 is the overall Panel 25 response rate through Round 5) plus 0.31 times 24.3 (24.3 is the overall Panel 26 response rate through Round 3.)

The overall response rate of 21.8 percent for 2021 is lower than that for 2020 (27.6 percent), reflecting the continued impact of the pandemic on data collection efforts.

3.3.2 Panel 26 Response Rates

For MEPS Panel 26 Round 1, 9,510 households were fielded in 2021 (Row C of Table 3.1), a randomly selected subsample of the households responding to the 2020 NHIS.

Table 3.1 shows the number of RUs eligible for interviewing in each Round of Panel 26 as well as the number of RUs completing the MEPS interview. Computing the individual round “conditional” response rates as described in Section 3.3.1 and then taking the product of these three response rates and the factor 60.6 (the percentage of the NHIS sampled households characterized as responding at the time of sample selection of households for MEPS) yields an overall response rate of 24.3 percent for Panel 26 through Round 3.

3.3.3 Panel 25 Response Rates

A total of 9,888 households were fielded in 2020 for MEPS Panel 25 (as indicated in Row C of Table 3.1), a randomly selected subsample of the households responding to the 2019 National Health Interview Survey (NHIS).

Table 3.1 shows the number of RUs eligible for interviewing and the number completing the interview for all five rounds of Panel 25. The overall response rate for Panel 25 was computed in

a similar fashion to that of Panel 26 but covering all five rounds of MEPS interviewing as well as the factor representing the percentage of NHIS sampled households eligible for MEPS. The overall response rate for Panel 25 through Round 5 is 19.6 percent.

3.3.4 Panel 24 Response Rates

A total of 9,684 households were fielded in 2019 for MEPS Panel 24 (as indicated in Row C of Table 3.1), a randomly selected subsample of the households responding to the 2018 National Health Interview Survey (NHIS).

Table 3.1 shows the number of RUs eligible for interviewing and the number completing the interview for all seven rounds of Panel 24. The overall response rate for Panel 24 was computed in a similar fashion to that of Panel 25 but covering all seven rounds of MEPS interviewing as well as the factor representing the percentage of NHIS sampled households eligible for MEPS. The overall response rate for Panel 24 through Round 7 is 20.7 percent.

3.3.5 Panel 23 Response Rates

A total of 9,694 households were fielded in 2018 for MEPS Panel 23 (as indicated in Row C of Table 3.1), a randomly selected subsample of the households responding to the 2017 National Health Interview Survey (NHIS).

Table 3.1 shows the number of RUs eligible for interviewing and the number completing the interview for all nine rounds of Panel 23. The overall response rate for Panel 23 was computed in a similar fashion to that of Panel 24 but covering all nine rounds of MEPS interviewing as well as the factor representing the percentage of NHIS sampled households eligible for MEPS. The overall response rate for Panel 23 through Round 9 is 22.0 percent.

3.3.6 Annual (Combined Panel) Response Rate

A combined panel response rate for the survey respondents in this data set is obtained by taking a weighted average of the panel-specific response rates. The Panel 23 response rate was weighted by a factor of 0.22, the Panel 24 response rate was weighted by a factor of 0.22, the Panel 25 response rate was weighted by a factor of 0.25, and the Panel 26 response rate was weighted by a factor of 0.31, reflecting approximately the distribution of the overall sample between the four panels. The resulting combined response rate for the combined panels was computed as $(0.22 \times 22.0) + (0.22 \times 20.7) + (0.25 \times 19.6) + (0.31 \times 24.3)$ or 21.8 percent (as shown in Table 3.1).

3.3.7 Oversampling

Oversampling is a feature of the MEPS sample design, helping to increase the precision of estimates for some subgroups of interest. Before going into details related to MEPS, the concept of oversampling will be discussed.

In a sample where all persons in a population are selected with the same probability and survey coverage of the population is high, the sample distribution is expected to be proportionate to the population distribution. For example, if Hispanics represent 15 percent of the general population, one would expect roughly 15 percent of the persons sampled to be Hispanic. However, in order to improve the precision of estimates for specific subgroups of a population, one might decide to select samples from those subgroups at higher rates than the remainder of the population. Thus, one might select Hispanics at twice the rate (i.e., at double the probability) of persons not oversampled. As a result, an oversampled subgroup comprises a higher proportion of the sample than it represents in the general population. Sample weights ensure that population estimates are not distorted by a disproportionate contribution from oversampled subgroups. Base sample weights for oversampled groups will be smaller than for the portion of the population not oversampled. For example, if a subgroup is sampled at roughly twice the rate of sample selection for the remainder of the population not oversampled, members of the oversampled subgroup will receive base or initial sample weights (prior to nonresponse or poststratification adjustments) that are roughly half the size of the group not oversampled.

As mentioned above, oversampling is implemented to increase the sample sizes and thus improve the precision of survey estimates for particular subgroups of the population. The “cost” of oversampling is that the precision of estimates for the general population and subgroups not oversampled will be reduced to some extent compared to the precision one could have achieved if the same overall sample size were selected without any oversampling.

The NHIS no longer oversamples households with members who are Asian, Black, or Hispanic. Nevertheless, these minority groups are still of analytic interest for MEPS. As a result for Panels 23, 24, and 25, all households in the Asian, Hispanic, and Black domains were sampled with certainty (i.e., all households assigned to those domains were included in the MEPS). In addition, all households in Panel 23 who had a member who was a veteran were also selected with certainty. Among all remaining households for Panel 23, the “Other, complete” domain was sampled at a rate of about 69 percent while the “Other, partial complete” domain was sampled at a rate of about 43 percent. For Panel 24, the corresponding sampling rates for the “Other, complete” domain and the “Other, partial complete” domain were about 79 percent and 50 percent, respectively. For Panel 25, the corresponding sampling rates for the “Other, complete” domain and the “Other, partial complete” domain were about 77 percent and 50 percent, respectively. The somewhat lower sampling rates for Panel 23 in the two “Other” domains arose due to the oversampling of veterans in that panel. With a specified overall sample size of 9,700 fewer were needed from those assigned to the “Other” domains in that panel.

Within the “noncertainty” strata (the “Other” domains) for the three panels, responding NHIS households were selected for MEPS using a systematic sample selection procedure from among those eligible. The selection of the households was with probability proportionate to size (pps) where the size measure was the inverse of the NHIS initial probability of selection. The pps sampling was undertaken to help reduce the variability in the MEPS weights incurred due to the variability of the NHIS sampling rates.

As discussed in Section 3.1.4, the Panel 26 sample focused on oversampling the “middle-sized” states rather than Asians, Blacks, and Hispanics.

A note with respect to the interpretation of MEPS response rates, which are unweighted. Typically, sample allocations across sample domains change from one MEPS panel to another. The sample domains used may also vary by panel as is the case for Panel 23 versus Panel 24, Panel 25, and Panel 26. When one compares unweighted measures (e.g., response rates) between panels and years, one should take into account such differences. Suppose, for example, members of one domain have a lower propensity to respond than those of another domain. Then, if that domain has been allocated a higher proportion of the sample, the corresponding panel may have a lower unweighted response rate simply because of the differences in sample allocation.

3.4 Background on Person-Level Estimation Using this MEPS Public Use Release

3.4.1 Requirements to Receive a Person-Level Weight

There is a single full year person-level weight (PERWT21F) assigned to each record for each key, in-scope person who responded to MEPS for the full period of time that they were in-scope during 2021. A key person is either a member of a responding NHIS household at the time of interview or joined a family associated with such a household after being out-of-scope at the time of the NHIS (the latter circumstance includes newborns as well as those returning from military service, an institution, or residence in a foreign country). A person is in-scope whenever they are a member of the civilian noninstitutionalized portion of the U.S. population.

3.4.2 Details on Person-Level Weights Construction

The person-level weight PERWT21F was developed in several stages. Initially, weights were developed for the Full Year 2021 Population Characteristics PUF, released a number of months earlier. Preliminary person-level weights for Panel 23, Panel 24, Panel 25, and Panel 26 were first created separately. The weighting process for each panel included adjustments for nonresponse over time and a calibration to the same independent population figures. Only those who were in-scope on December 31, 2021 were raked to Current Population Survey (CPS) control totals.

The calibration was undertaken for each panel separately by raking the nonresponse adjusted weights to CPS population estimates based on six variables. The six variables used in the establishment of the person-level control figures were: education of the reference person (no degree, high school/GED no college or some college, Bachelor's degree or higher); census region (Northeast, Midwest, South, West); MSA status (MSA, non-MSA); race/ethnicity (Hispanic; Black, non-Hispanic; Asian, non-Hispanic; and other); sex; and age. (It may be noted that, for confidentiality reasons, the MSA status variables are no longer released for public use. This started with the Full-Year 2013 Person-Level Use PUF.)

A weight for the full sample was created by applying compositing factors to the weights associated with each panel where the factors are all greater than 0 and sum to 1. The compositing factors for these weights reflect the relative "effective sample size" of each panel, representing

the inverse of the relative amount of variability in the individual panel estimates attributable to the variability of the sample weights and sample size. Using these factors helps limit the variability of the estimates obtained using the composited weights across the four samples pooled.

The effective sample sizes were computed for each panel by dividing the sample size by the design effect associated with the variability of the nonresponse-adjusted person weights in that panel (i.e., prior to raking the weights of a panel) across the person-level respondents in the panel. The relative effective sample size was then computed by taking the ratio of the effective sample size for a panel to the sum of the effective sample sizes across the four panels.

Using the relative nominal sample sizes (the proportions that the number of respondents in a panel represent among the total number of respondents in the four panels) as compositing factors has worked well for MEPS in previous years. However, using the relative effective sample size for 2021 data increases the effectiveness of the compositing factors to limit variance to some extent in a year where the panel weights are more variable due to the higher than usual nonresponse. This approach is planned for MEPS in future years as well.

In terms of the actual details, the individual panel weights from Panel 23 were multiplied by the factor .22, each weight from Panel 24 was multiplied by the factor .22, each weight from Panel 25 by the factor .25, and each weight from Panel 26 by the factor .31. Once the compositing factors were applied to produce a full-sample weight, a final raking was undertaken, based on the same six variables used in the raking of the individual panel weights. This weight served as the person-level weight appearing on the Full Year 2021 Population Characteristics PUF.

The standard approach for establishing the Full Year 2021 Consolidated PUF weight from the Full Year 2021 Population Characteristics PUF is as follows. When MEPS poverty status information derived from income variables becomes available, a final raking is undertaken. The full sample weight appearing on the Population Characteristics PUF for a given year is re-raked, replacing educational attainment with poverty status while retaining the other five raking variables previously indicated. Specifically, control totals based on CPS estimates of poverty status (five categories: below poverty, from 100 to 125 percent of poverty, from 125 to 200 percent of poverty, from 200 to 400 percent of poverty, at least 400 percent of poverty) as well as age, race/ethnicity, sex, region, and MSA status were used in the weight calibration process. Only those in scope on December 31, 2021 were included in the raking process.

3.4.3 MEPS Panel 23 Weight Development Process

The person-level weight for MEPS Panel 23 was developed using the 2020 full-year weight for an individual as the initially assigned weight for 2020 survey participants present in 2021. For key, in-scope members who joined an RU some time in 2021 after being out-of-scope in 2020, the initially assigned person-level weight was the corresponding 2020 family weight. The weighting process included an adjustment for person-level nonresponse over Rounds 8 and 9 as well as raking to population control figures for December 2021 for key, responding persons in-scope on December 31, 2021. These control totals were derived by scaling back the population distribution obtained from the March 2022 CPS to reflect the December 31, 2021 estimated

population total (estimated based on Census projections for January 1, 2022). Variables used for person-level raking included: education of the reference person (three categories: no degree; high school/GED only or some college; Bachelor's or higher degree); Census region (Northeast, Midwest, South, West); MSA status (MSA, non-MSA); race/ethnicity (Hispanic; Black, non-Hispanic; Asian, non-Hispanic; and other); sex; and age. (It may be noted that for confidentiality reasons, the MSA status variables are no longer released for public use.) The final weight for key, responding persons who were not in-scope on December 31, 2021 but were in-scope earlier in the year was the nonresponse-adjusted person weight without raking. ”

The 2020 full-year weight used as the base weight for Panel 23 was derived from the 2018 MEPS Round 1 weight and reflected adjustment for nonresponse over the remaining data collection rounds in 2018, 2019, and 2020 as well as raking to the December 2018, December 2019, and December 2020 population control figures.

3.4.4 MEPS Panel 24 Weight Development Process

The person-level weight for MEPS Panel 24 was developed using the 2020 full-year weight for an individual as a “base” weight for survey participants present in 2021. For key, in-scope members who joined an RU some time in 2021 after being out-of-scope in 2020, the initially assigned person-level weight was the corresponding 2020 family weight. The weighting process also included an adjustment for person-level nonresponse over Rounds 6 and 7 as well as raking to the same population control figures for December 2021 used for the MEPS Panel 23 weights for key, responding persons in-scope on December 31, 2021. The same six variables employed for Panel 23 raking (education level, census region, MSA status, race/ethnicity, sex, and age) were also used for Panel 24 raking. Similar to Panel 23, the Panel 24 final weight for key, responding persons not in-scope on December 31, 2021 but in-scope earlier in the year was the nonresponse-adjusted person weight without raking.

Note that the 2020 full-year weight that was used as the base weight for Panel 24 was derived using the 2019 MEPS Round 1 weight and reflected adjustment for nonresponse over the remaining data collection rounds in 2019 and 2020 as well as raking to the December 2019 and December 2020 population control figures.

3.4.5 MEPS Panel 25 Weight Development Process

The person-level weight for MEPS Panel 25 was developed using the 2020 full-year weight for an individual as a “base” weight for survey participants present in 2021.

For key, in-scope members who joined an RU sometime in 2021 after being out-of-scope in 2020, the initially assigned person-level weight was the corresponding 2020 family weight. The weighting process also included an adjustment for person-level nonresponse over Rounds 4 and 5 as well as raking to the same population control figures for December 2021 used for the MEPS Panels 23 and 24 weights for key, responding persons in-scope on December 31, 2021. The same six variables employed for Panels 23 and 24 raking (education level, census region, MSA status, race/ethnicity, sex, and age) were also used for Panel 25 raking. Similar to Panels 23 and 24, the

Panel 25 final weight for key, responding persons not in-scope on December 31, 2021 but in-scope earlier in the year was the nonresponse-adjusted person weight without raking.

Note that the 2020 full-year weight that was used as the base weight for Panel 25 was derived using the 2020 MEPS Round 1 weight and reflected adjustment for nonresponse over the remaining data collection rounds in 2020 as well as raking to the December 2020 population control figures.

3.4.6 MEPS Panel 26 Weight Development Process

The person-level weight for MEPS Panel 26 was developed using the 2021 MEPS Round 1 person-level weight as a “base” weight. The MEPS Round 1 weights incorporated the following components: the original household probability of selection for the NHIS adjusted to reflect the NHIS subsample reserved for MEPS and an adjustment for NHIS nonresponse, the probability of selection for MEPS from NHIS, an adjustment for nonresponse at the dwelling unit level for Round 1, and poststratification to control figures at the person level obtained from the March CPS of 2021. For key, in-scope members who joined an RU after Round 1, the Round 1 DU weight served as a “base” weight.

The weighting process also included an adjustment for nonresponse over the remaining data collection rounds in 2021 as well as raking to the same population control figures for December 2021 used for the MEPS Panel 23, Panel 24, and Panel 25 weights for key, responding persons in-scope on December 31, 2021. The same six variables employed for Panel 23, Panel 24, and Panel 25 raking (education level of the reference person, census region, MSA status, race/ethnicity, sex, and age) were also used for Panel 26 raking. Similar to Panel 23, Panel 24, and Panel 25, the Panel 26 final weight for key, responding persons who were not in-scope on December 31, 2021 but were in-scope earlier in the year was the nonresponse-adjusted person weight without raking.

3.4.7 The Final Person-Level Weight for 2021

As described in Section 3.4.2, person-level weights from the four panels were then composited using factors identified in 3.4.2 and, for those persons inscope on December 31, raked to CPS based control totals reflecting poverty status as well as census region, MSA status, race/ethnicity, sex, and age. In addition, the composite weights of two groups of persons who were out-of-scope on December 31, 2021 underwent weight calibration to adjust for expected undercoverage. Specifically, the weights of those who were inscope sometime during the year, out-of-scope on December 31, and entered a nursing home during the year were poststratified to an estimate of the number of persons who were residents of Medicare- and Medicaid-certified nursing homes for part of the year (approximately 3-9 months) during 2014. This estimate was developed from data on the Minimum Data Set (MDS) of the Center for Medicare and Medicaid Services (CMS). The weights of persons who died while inscope during 2021 were poststratified to corresponding estimates derived using data obtained from the Centers for Disease Control and Prevention (CDC), National Center for Health Statistics (NCHS), Underlying Cause of Death, 2018-2021 on [CDC WONDER Online Database](#), released in 2023, the latest available data at the time. Separate

decendent control totals were developed for the “65 and older” and “under 65” civilian noninstitutionalized populations.

The sum of the person-level weights across all persons assigned a positive person-level weight, (i.e., for the civilian, noninstitutionalized or in-scope population over the course of the year (based on PERWT21F>0) is 331,249,393 (see Table 3.2). The corresponding total for the population that was inscope on December 31, 2021 is 327,209,772.

Table 3.2. Number of person-level respondents and corresponding population estimates for the 2021 Full Year Consolidated File

Populations of Interest	Panel 23	Panel 24	Panel 25	Panel 26	Combined	Population estimate (weighted total of combined samples)
Civilian, Noninstitutionalized Population over the course of 2021	6,674	6,441	5,990	8,227	27,332	331,249,393
Civilian, Noninstitutionalized Population on December 31, 2021	6,596	6,370	5,934	8,118	27,018	327,209,772

3.4.8 A Note on MEPS Population Estimates

Beginning with the 2021 Full Year data, MEPS transitioned to 2020 census-based population estimates from the CPS for poststratification and raking. CPS estimates began reflecting 2020 census-based data in 2022, and the March 2022 CPS data serve as the basis for the 2021 MEPS weight calibration efforts.

Use of the updated population controls will have a noticeable effect on estimated totals for some population subgroups. The article [Adjustments to Household Survey Population Estimates in January 2023](#) compares some 2022 CPS estimates for those aged 16 and older “as published” with those that would have been generated had the updated population controls been used. The more notable increase among the subgroups was for Asians (440 thousand more, a 2.6 percent increase). Corresponding changes were thus anticipated for MEPS full year data beginning with the 2021 MEPS PUF.

3.5 Coverage

The target population associated with this MEPS database is the 2021 U.S. civilian, noninstitutionalized population. However, the MEPS sampled households are a subsample of the NHIS households interviewed in 2017 (Panel 23), 2018 (Panel 24), 2019 (Panel 25), and 2020

(Panel 26). New households created after the NHIS interviews for the respective panels and consisting exclusively of persons who entered the target population after 2017 (Panel 23), after 2018 (Panel 24), after 2019 (Panel 25), or after 2020 (Panel 26) are not covered by MEPS. Neither are previously out-of-scope persons who join an existing household but are unrelated to the current household residents. Persons not covered by a given MEPS panel thus include some members of the following groups: immigrants, persons leaving the military, U.S. citizens returning from residence in another country, and persons leaving institutions. Those not covered represent a small proportion of the MEPS target population.

3.6 Background on Family-Level Estimation Using This MEPS Public Use File

3.6.1 Overview

There are two family weight variables provided in this release: FAMWT21F and FAMWT21C. FAMWT21F can be used to make estimates for the cross-section of families in the U.S. civilian noninstitutionalized population on December 31, 2021 where families are identified based on the MEPS definition of a family unit. Estimates can include MEPS families that existed at some time during 2021 but whose members became out-of-scope prior to the end of the year (e.g., all family members moved out of the country, died, etc.) as well as MEPS families in existence on December 31, 2021. FAMWT21C can be used to make estimates for the cross-section of families in the U.S. civilian, noninstitutionalized population on December 31, 2021 where families are identified based on the CPS definition of a family unit. It may be noted that married couples, regardless of gender, are assigned to the same family unit for the CPS. This represents a definitional change in family structure for the CPS with the 2021 CPS ASEC file, associated specifically with March, 2021. It is reflected in MEPS starting with the 2019 Full Year Consolidated PUF.

3.6.2 Definition of “Family” for Estimation Purposes

A MEPS family generally consists of two or more persons living together in the same household who are related by blood, marriage, or adoption. MEPS also defines as a family unmarried persons living together who consider themselves a family unit (these are not families under the CPS definition). Single people who live with neither a relative nor a person identified as a “significant other” have also been assigned a family ID value and a family-level weight. Thus, they can be included or excluded from family-level estimates, as desired. Relatives identified as usual residents of the household who were not present at the time of the interview, such as college students living away from their parents’ home during the school year, were considered as members of the family that identified them.

To make estimates at the family level, it is necessary to prepare a family-level file containing one record per family (see instructions below), family-level summary characteristics, and the family-level weight variable (FAMWT21F or FAMWT21C). Each MEPS family unit is uniquely identified by the combination of the variables DUID and FAMIDYR while each CPS family unit

is uniquely identified by the combination of the variables DUID and CPSFAMID. Only persons with positive nonzero family weight values are candidates for inclusion in family estimates.

Two sets of families for whom estimates can be obtained are defined in Table 3.3 below (along with respective sample sizes). Persons with FMRS1231=1 were a member of a MEPS family on 12/31/21. The more expansive definition of families (second row in Table 3.3) includes families and members of families who were not inscope at the end of the year. While MEPS includes individual persons as family units (about one-third of all units), analysts may restrict their analyses to families with two or more members using the family size variables shown in Table 3.3 (for example, to limit consideration to the cross-section of families with two or more members on December 31, 2021, analyze only families where FAMS1231 is 2 or more). Estimates can also be made for the cross-section of CPS families on December 31, 2021 based on the 13,461 sample CPS families in this data file.

Table 3.3. Identifying MEPS Families and Corresponding Sample Sizes

Population of Interest	Cases to Include	Sample Size (Includes single person units)	Family Size Variable
Cross-section of Families in the Civilian Noninstitutionalized Population on 12/31/21	FAMWT21F>0 & FMRS1231=1	12,910	FAMS1231
Families in the Civilian Noninstitutionalized Population on 12/31/21 plus families and members of families in existence earlier in 2021 who were not part of the civilian noninstitutionalized population on 12/31/21	FAMWT21F>0	13,014	FAMSZEYR

3.6.3 Instructions to Create Family Estimates

The following is a summary of the steps and the variables to be used for family-level estimation based on the MEPS definition of families.

- Restrict the records of interest to those where FAMWT21F>0
- Concatenate the variables DUID and FAMIDYR into a new variable (e.g., DUIDFAMY).
- To create a family-level file, sort by DUIDFAMY and then subset to one record per DUIDFAMY value by retaining only the reference person record (FAMRFPYR=1) for each value of DUIDFAMY. Some family-level measures needed for analytic purposes (e.g., means or totals) can be obtained after aggregating person-level information across all members of a family. For other types of measures, analysts frequently use the characteristics of the reference person to characterize their family unit (e.g., the race/ethnicity, marital status, or age of the reference person).

- Apply the weight FAMWT21F to the analytic variable(s) of interest to obtain national MEPS family estimates.

It should be noted that the MEPS families defined above include members who were out-of-scope on December 31, 2021 although they were members of the family immediately prior to going out of scope for the remainder of the year. If an analyst wishes to restrict MEPS family members to those who were a family member on December 31, 2021, the analyst should restrict family members to those with FMRS1231=1.

The following is a summary of the steps and the variables to be used for family-level estimation based on the CPS definition of families.

- Restrict the records of interest to those where FAMWT21C>0
- Concatenate the variables DUID and CPSFAMID into a new variable (e.g., DUIDFAMC).
- To create a family-level file, sort by DUIDFAMC and then subset to one record per DUIDFAMC value by retaining only the reference person record (FCRP1231=1) for each value of DUIDFAMC. Some family-level measures needed for analytic purposes (e.g., means or totals) can be obtained after aggregating person-level information across all members of a family. For other types of measures, analysts frequently use the characteristics of the reference person to characterize their family unit (e.g., the race/ethnicity, marital status, or age of the reference person). (Note that to be strictly comparable to the CPS definition of families, only those with two or more family members should be included in analyses.)
- Apply the weight FAMWT21C to the analytic variable(s) of interest to obtain national CPS family estimates.

It should be noted that these CPS-families consist solely of those who were family members on December 31, 2021.

3.6.4 Details on Family Weight Construction and Estimated Number of Families

Because health care related decisions are influenced by a family's economic status, poverty status is incorporated into the poststratification component of the weighting process. However, poverty status is defined based on the CPS definition of a family, which differs from the MEPS family definition in that unmarried partners living together are considered separate family units. Since data are collected in MEPS family units (RUs), prior to poststratification MEPS families in existence on December 31, 2021 containing unmarried partners living together were partitioned into units that correspond to CPS families (families with no unmarried partners are defined as family units in both MEPS and CPS).

The process of calibrating the family weights to achieve consistency with CPS control figures was carried out in several steps. First, all CPS-like family units were assigned an initial family-level weight based on the person-level weight (PERWT21F) of the family reference person

(FAMRFPYR=1) of the MEPS family with which they were associated. These CPS family-level weights (FAMWT21C) were obtained by raking to population control figures derived from CPS estimates for December 2021 (derived by projecting the family population totals for the March 2021 CPS forward to reflect December 31, 2021). In addition to poverty status, the calibration process for the family-level weights incorporated the following variables: Census region; MSA status; race/ethnicity of reference person (Hispanic, Black but non-Hispanic, Asian, and other); family type (reference person married, living with spouse; male reference person, unmarried or spouse not present; female reference person, unmarried or spouse not present); age of reference person; and family size on December 31, 2021. The family-level weight variable for MEPS families (FAMWT21F) was then constructed by putting MEPS families that consisted of more than one CPS-like family back together and assigning the MEPS family-level weight based on the CPS family weight of the MEPS family reference person.

The weighted population estimate for CPS families on December 31, 2021 based on 13,461 CPS families in the sample is 143,980,355. Overall, the weighted population estimate for the 12,910 MEPS family units containing at least one member of the U.S. civilian, noninstitutionalized population on December 31, 2021 (those families whose members have FAMWT21F>0 and FMRS1231=1) is 137,896,483. The inclusion of families whose members left the in-scope population prior to December 31, 2021 increases the estimated total number of families represented by the 13,014 MEPS responding families (whose members have FAMWT21F>0) to 142,463,160.

Table 3.4. Families with a family weight >0 for the 2021 Full Year Consolidated Data File

	Panel 23	Panel 24	Panel 25	Panel 26	Combined	Population estimate (weighted total of combined samples)
Number	3,227	3,177	2,874	3,736	13,014	142,463,160

3.7 Analysis Using Health Insurance Eligibility Units

To construct a weight for use in analysis using Health Insurance Eligibility Units, as identified by the variable HIEUIDX:

1. Identify the HIEU head by your analytic intent, i.e. if only studying health insurance unit with female heads of households, choose the female adult as head of household.
2. If the weight of the HIEU head is non-zero, use the weight of the HIEU head for all members of that HIEU; or

If the weight of the HIEU head is zero, delete the case.

3.8 Weights and Response Rates for the Self-Administered Questionnaire

For analytic purposes, a single person-level weight variable, SAQWT21F, has been provided for use with the data obtained from the Self-Administered Questionnaire (SAQ). This questionnaire was administered in Panel 26 Round 2, Panel 25 Round 4, Panel 24 Round 6, and Panel 23 Round 8 and was to be completed by each adult (person aged 18 or older) in the family. Thus, the target population for the SAQ is adults in the civilian, noninstitutionalized population at the time data were collected for Rounds 2/4/6/8 (generally speaking, the fall of the year in question).

The final full-year person-level SAQ weight for 2021 was constructed as follows with only those with a 2021 full year person weight ($PERWT21F > 0$) eligible to receive the 2021 SAQ weight. The weighting process was similar to that of the full sample person-level weights: nonresponse adjustments for the weights for each panel separately; raking to CPS control totals; compositing the weights from the three panels; and finally re-raking of the composited weights.

Variables used in the nonresponse adjustment process were region, MSA status, family size, marital status, level of education, health status, health insurance status, age, sex, and race/ethnicity. The weights were raked to Current Population Survey (CPS) estimates corresponding to December 2021 (the same source of control figures used for the full year person weights). The variables used to form control figures (education of the reference person, region, MSA status, age, sex, and race/ethnicity) are the same variables that were used for the full year person weights. The only difference was that age categories were developed after excluding ages under 18, since only adults were eligible for the SAQ.

The final 2021 SAQ weight for this consolidated data file was then obtained by raking the preliminary weight to CPS estimates that were based on poverty status (replacing education of the reference person) as well as the aforementioned variables. This final weight was assigned the variable name SAQWT21F.

In all, there were 15,422 persons assigned an SAQ weight with the sum of the weights being 254,207,493 (an estimate of the civilian, noninstitutionalized population aged 18 or older at the time the SAQ was administered).

The Panel 23 unweighted response rate for the 2021 SAQ was 63.6 percent, the Panel 24 unweighted response rate for the 2021 SAQ was 60.4 percent, the Panel 25 unweighted response rate for the 2021 SAQ was 65.1 percent, and the Panel 26 unweighted response rate for the 2021 SAQ was 64.9 percent. Pooled unweighted response rates for the survey respondents have been computed by taking a weighted average of the panel-specific response rates, where the weights were the same factors used for the weight compositing as described in Section 3.4.2 (a value of 0.22 was associated with Panel 23, a value of 0.22 was associated with Panel 24, a value of 0.25 was associated with Panel 25, and a value of 0.31 was associated with Panel 26). The pooled unweighted response rate for the combined panels for the 2021 SAQ is 63.7 percent.

3.9 Weights and Response Rates for the Diabetes Care Survey

A person-level weight, DIABW21F, was developed for use with the data obtained from the Diabetes Care Survey (DCS). This weight was assigned to each person aged 18 or over with an SAQ weight who completed the DCS and self-reported as having diabetes. (Although diabetes diagnosis is now asked of all ages, the DCS is only given to participants if they are 18 years of age or older.) The general weighting process was to assign the SAQ expenditure weight as the initial weight to each individual eligible for a DCS weight. This weight was adjusted to compensate for RU level nonresponse to the question as to whether or not each RU member had diabetes and then for nonresponse among those receiving the DCS questionnaire.

Prior to Panel 12, the identification of people eligible to receive the DCS questionnaire was focused on the Rounds 3/5 interview. During the Rounds 3/5 regular MEPS interview, each RU respondent was asked to complete a “conditions” question to identify all current/deceased/institutionalized RU members of any age who had been diagnosed with diabetes. Each RU member who was identified as having diabetes by the RU respondent was then eligible to receive the DCS questionnaire. To determine which DCS respondents actually had diabetes (and thus were members of the target population), each DCS respondent was asked if s/he was told by a physician that s/he had diabetes. While the DCS questionnaire was distributed to persons under the age of 18, the constructed DCS variables released in the person-level PUF applied only to adults. Beginning in Panel 12, a different screening process has been employed to identify those eligible to receive the DCS questionnaire. This process, described next, involves asking screener questions in each round, with the group of persons about whom these questions are asked varying from round to round.

In Round 1, the RU respondent is asked to identify all RU members (including those who went out of scope unless they died prior to the date of interview) with diabetes. In Rounds 2/4/6/8, the same screening information is gathered but only for new RU members (as long as they did not die during the round). In Round 3, Round 5, and Round 7 the screening questions are asked of the RU respondent for all RU members who were: (a) inscope sometime during the round but had not died prior to the date of interview; and (b) had not been identified as having diabetes in a previous round (this includes people with missing data, classified as not having diabetes in all previous rounds of MEPS, and all new members of the RU in Round 3, Round 5, or Round 7). Any RU member 18 years or older who has been identified by the RU respondent as having diabetes at any time during MEPS will be asked to complete a DCS questionnaire. This process has been designed to help ensure that all RU members who were 18 years or older with diabetes will be given a DCS questionnaire to complete.

In all, 1,640 people were assigned a DCS weight (DIABW21F>0). The sum of the DCS weights is 27,972,020, an estimate of the adult population self-reporting as having been diagnosed with diabetes based on the two-step process described above.

The Panel 23 unweighted response rate for the 2021 DCS was 54.9 percent. The Panel 24 unweighted response rate for the 2021 DCS was 54.2 percent. The Panel 25 unweighted response rate for the 2021 DCS was 55.3 percent. The Panel 26 unweighted response rate for the 2021 DCS was 57.7 percent. The pooled unweighted response rate, calculated based on applying the

compositing factors described in Section 3.4.2, for the combined panels for the DCS is 55.7 percent.

3.10 Weights and Response Rates for the Social Determinants of Health Survey

For analytic purposes, a single person-level weight variable, SDOHWT21F, has been provided for use with the data obtained from the Social Determinants of Health (SDOH) questionnaire. This questionnaire was administered in Panel 26, Round 1, Panel 25 Round 3, Panel 24 Round 5, and Panel 23 Round 7. These were the rounds associated with the initial fielding of each panel in 2021. The questionnaire was to be completed by each adult (person aged 18 or older) in the family. Thus, the target population for the SDOH is adults in the civilian, noninstitutionalized population at the time data were collected for Rounds 1/3/5/7 (generally speaking, the late winter/spring of 2021).

The final full-year person-level SDOH weight for 2021 was constructed as follows with only those with a 2021 full year person weight (PERWT21P>0) eligible to receive the 2021 SDOH weight. The weighting process for the SDOH weight initiated with the 2021 full year person weights of each panel. The full sample person-level weights were adjusted to account for nonresponse to the SDOH for each panel separately. A raking process similar to the full sample weight were then implemented for each panel. The final steps were to composite the weights from the four panels and re-rake the composited weights.

March 2021 CPS control totals were used to rake each panel separately.

Variables used in the nonresponse adjustment process were region, MSA status, family size, marital status, level of education, employment status of the reference person, health status, health insurance status, age, sex, and race/ethnicity. The March 2021 CPS control totals were formed based on the following variables: educational attainment of the reference person, region, MSA status, age, sex, race/ethnicity, and poverty status. The age categories were developed after excluding ages under 18 since only adults were eligible for the SDOH.

This final weight was assigned the variable name SDOHWT21F.

In all, there were 18,400 persons assigned an SDOH weight with the sum of the weights being 252,273,615 (an estimate of the civilian, noninstitutionalized population aged 18 or older at the time the SDOH was administered).

The unweighted response rates for the 2021 SDOH for each panel were: Panel 23, 79.7 percent; Panel 24, 75.5 percent; Panel 25, 76.8 percent; and Panel 26, 68.5. The pooled unweighted response rate for the SDOH survey respondents has been computed by taking a weighted average of the panel-specific response rates, where the weights were the same factors used for the weight compositing as described in Section 3.3 (these values are: Panel 23, 0.22; Panel 24, 0.22; Panel 25, 0.25; and Panel 26, 0.31). The pooled unweighted response rate for the combined panels for the 2021 SDOH is 74.5 percent.

3.11 Variance Estimation

To obtain estimates of variability (such as the standard error of sample estimates or corresponding confidence intervals) for MEPS estimates, analysts need to take into account the complex sample design of MEPS for both person-level and family-level analyses. Several methodologies have been developed for estimating standard errors for surveys with a complex sample design, including the Taylor-series linearization method, balanced repeated replication, and jackknife replication. Various software packages provide analysts with the capability of implementing these methodologies. MEPS analysts most commonly use the Taylor Series approach. Although this data file does not contain replicate weights, the capability of employing replicate weights constructed using the Balanced Repeated Replication (BRR) methodology is also provided if needed to develop variances for more complex estimators (see Section 3.11.2).

3.11.1 Taylor-series Linearization Method

The variables needed to calculate appropriate standard errors based on the Taylor-series linearization method are included on this and all other MEPS public use files. Software packages that permit the use of the Taylor-series linearization method include SUDAAN, R, Stata, SAS (version 8.2 and higher), and SPSS (version 12.0 and higher). For complete information on the capabilities of a package, analysts should refer to the corresponding software user documentation.

Using the Taylor-series linearization method, variance estimation strata and the variance estimation PSUs within these strata must be specified. The variables VARSTR and VARPSU on this MEPS data file serve to identify the sampling strata and primary sampling units required by the variance estimation programs. Specifying a “with replacement” design in one of the previously mentioned computer software packages will provide estimated standard errors appropriate for assessing the variability of MEPS survey estimates. It should be noted that the number of degrees of freedom associated with estimates of variability indicated by such a package may not appropriately reflect the number available. For variables of interest distributed throughout the country (and thus the MEPS sample PSUs), one can generally expect to have at least 100 degrees of freedom associated with the estimated standard errors for national estimates based on this MEPS database.

Prior to 2002, MEPS variance strata and PSUs were developed independently from year to year, and the last two characters of the strata and PSU variable names denoted the year. Beginning with the 2002 Point-in-Time PUF, the approach changed with the intention that variance strata and PSUs would be developed to be compatible with all future PUFs until the NHIS design changed. Thus, when pooling data across years 2002 through the Panel 11 component of the 2007 files, the variance strata and PSU variables provided can be used without modification for variance estimation purposes for estimates covering multiple years of data. There are 203 variance estimation strata, each stratum with either two or three variance estimation PSUs.

From Panel 12 of the 2007 files, a new set of variance strata and PSUs were developed because of the introduction of a new NHIS design. There are 165 variance strata with either two or three variance estimation PSUs per stratum, starting from Panel 12. Therefore, there are a total of 368

(203+165) variance strata in the 2007 Full Year file as it consists of two panels that were selected under two independent NHIS sample designs. Since both MEPS panels in the Full Year files from 2008 through 2016 were based on the next NHIS design, there are only 165 variance strata. These variance strata (VARSTR values) have been numbered from 1001 to 1165 so that they can be readily distinguished from those developed under the former NHIS sample design if data are pooled for several years.

The NHIS sample design was changed again in 2016, effectively changing the MEPS design beginning with calendar year 2017. From Panel 22 of the 2017 files, a new set of variance strata and PSUs were developed. There are 117 variance strata with either two or three variance estimation PSUs per stratum. Therefore, there are a total of 282 (165+117) variance strata in the 2017 Full Year file as it consists of two panels that were selected under two independent NHIS sample designs. To make the pooling of data across multiple years of MEPS more straightforward, the numbering system for the variance strata has changed. Those strata associated with the new design were numbered from 2001 to 2117.

However, the NHIS sample design was further modified in 2018. With the modification in the 2018 NHIS sample design, the MEPS variance structure for the 2019 Full Year file was also modified, reducing the number of variance strata to 105. Consistency was maintained with the prior structure in that the 2019 Full Year file variance strata were also numbered within the range of values from 2001-2117, although there are now gaps in the values assigned within this range. Due to the modification, each stratum could contain up to five variance estimation PSUs.

For Panel 26 in the 2021 Full Year file, additional NHIS sample was used for MEPS to account for increasing nonresponse during the pandemic (as discussed in Section 3.1.4). The additional sample was assigned to the existing variance strata, so the 2021 Full Year file continues to have 105 variance strata, numbered 2001-2117, with a few gaps in the values in that range. In many cases, the additional sample was assigned to new variance estimation PSUs, so in the 2021 Full Year file, each stratum could contain up to eight variance estimation PSUs.

Some analysts may be interested in pooling data across multiple years of MEPS data. If pooling across years is to be undertaken, it should be noted that, to obtain appropriate standard errors when doing so, it is necessary to specify a common variance structure. Prior to 2002, each annual MEPS public use file was released with a variance structure unique to the particular MEPS sample in that year. Starting in 2002, the annual MEPS public use files were released with a common variance structure that allowed users to pool data from 2002 through 2018. However, with the need to modify the variance structure beginning with 2019, this can no longer be routinely done.

To ensure that variance strata are identified appropriately for variance estimation purposes when pooling MEPS data across several years, one can proceed as follows:

1. When pooling any years between 2002 through 2018, use the variance strata numbering as is.
2. When pooling (a) any year from 1996 to 2001 with any year from 2002 or later, or (b) the year 2019 and beyond with any earlier year, use the pooled linkage public use

file HC-036 that contains the proper variance structure. The HC-036 file is updated every year so that appropriate variance structures are available with pooled data. Further details on the HC-036 file can be found in the public use documentation of the HC-036 file.

3.11.2 Balanced Repeated Replication (BRR) Method

BRR replicate weights are not provided on this MEPS PUF for the purposes of variance estimation. However, a file containing a BRR replication structure is made available so that the users can form replicate weights, if desired, from the final MEPS weight to compute variances of MEPS estimates using either BRR or Fay's modified BRR (Fay, 1989) methods. The replicate weights are useful to compute variances of complex non-linear estimators for which a Taylor linear form is not easy to derive and not available in commonly used software. For instance, it is not possible to calculate the variances of a median or the ratio of two medians using the Taylor linearization method. For these types of estimators, users may calculate a variance using BRR or Fay's modified BRR methods. However, it should be noted that the replicate weights have been derived from the final weight through a shortcut approach. Specifically, the replicate weights are not computed starting with the base weight and all adjustments made in different stages of weighting are not applied independently in each replicate. Thus, the variances computed using this one-step BRR do not capture the effects of all weighting adjustments that would be captured in a set of fully developed BRR replicate weights. The Taylor Series approach does not fully capture the effects of the different weighting adjustments either.

The dataset HC-036BRR, MEPS 1996-2021 Replicates for Variance Estimation File, contains the information necessary to construct the BRR replicates. It contains a set of 128 flags (BRR1-BRR128) in the form of half sample indicators, each of which is coded 0 or 1 to indicate whether the person should or should not be included in that particular replicate. These flags can be used in conjunction with the full-year weight to construct the BRR replicate weights. For analysis of MEPS data pooled across years, the BRR replicates can be formed in the same way using the HC-036, MEPS 1996-2021 Pooled Linkage Variance Estimation File. For more information about creating BRR replicates, users can refer to the documentation for the [HC-036BRR pooled linkage file](#) on the AHRQ website.

3.12 Guidelines for Determining which Weight to Use for Analyses Involving Data/Variables from Multiple Sources and Supplements: MEPS 2021 Full-Year Use File

Which weight variable to use is decided based on a hierarchy.

For person-level analyses not involving variables from the SAQ, DCS, or SDOH, PERWT21F should always be used.

For person-level analysis involving variables from the SAQ but not the DCS, the weight variable SAQWT21F should be used. For example, if examining access to care or quality of care variables from the SAQ by socio-demographics, health status, or health insurance status,

SAQWT21F is the appropriate weight even though person-level socio-demographic, health status, and health insurance status variables are part of the core person-level questionnaire. Whenever data from the Diabetes Care Survey (DCS) are used, alone or in conjunction with data from other questionnaires, the weight variable DIABW21F should be used for those eligible to provide DCS data.

For person-level analysis involving variables from the SDOH, the weight variable SDOHWT21F should be used.

For all family-level analyses, FAMWT21F or FAMWT21C should be used.

3.13 Using MEPS Data for Trend Analysis

First, of course, we note that there are uncertainties associated with 2020 and 2021 data quality as discussed earlier in the data quality section (Section 3.1). Evaluations of important MEPS estimates suggest that they are of reasonable quality. Nevertheless, analysts are advised to exercise caution in interpreting these estimates, particularly in terms of trend analyses since access to health care was substantially affected by the COVID-19 pandemic as were related factors such as health insurance and employment status for many people.

MEPS began in 1996, and the utility of the survey for analyzing health care trends expands with each additional year of data; however, when examining trends over time using MEPS, the length of time being analyzed should be considered. In particular, large shifts in survey estimates over short periods of time (e.g., from one year to the next) that are statistically significant should be interpreted with caution unless they are attributable to known factors such as changes in public policy, economic conditions, or MEPS survey methodology.

With respect to methodological considerations, in 2013 MEPS introduced an effort focused on field procedure changes such as interviewer training to obtain more complete information about health care utilization from MEPS respondents with full implementation in 2014. This effort likely resulted in improved data quality and a reduction in underreporting starting in the second half of 2013 and throughout the 2014 full year files and have had some impact on analyses involving trends in utilization across years.

The changes in the NHIS sample design in 2016 and 2018 could also potentially affect trend analyses. The new NHIS sample design is based on more up-to-date information related to the distribution of housing units across the U.S. As a result, it can be expected to better cover the full U.S. civilian, noninstitutionalized population, the target population for MEPS, as well as many of its subpopulations. Better coverage of the target population helps to reduce the potential for bias in both NHIS and MEPS estimates.

Another change with the potential to affect trend analysis involved major modifications to the MEPS instrument design and data collection process, particularly in the events sections of the instrument. These were introduced in the Spring of 2018 and thus affected data beginning with Round 1 of Panel 23, Round 3 of Panel 22, and Round 5 of Panel 21. Since the Full Year 2017 PUFs were established from data collected in Rounds 1-3 of Panel 22 and Rounds 3-5 of Panel

21, they reflected two different instrument designs. In order to mitigate the effect of such differences within the same full year file, the Panel 22 Round 3 data and the Panel 21 Round 5 data were transformed to make them as consistent as possible with data collected under the previous design. The changes in the instrument were designed to make the data collection effort more efficient and easy to administer. In addition, expectations were that data on some items, such as those related to health care events, would be more complete with the potential of identifying more events. Increases in service use reported since the implementation of these changes are consistent with these expectations. ***Data users should be aware of possible impacts on the data and especially trend analysis for these data years due to the design transition.***

Process changes, such as data editing and imputation, may also affect trend analyses. For example, users should refer to Section 2.5.11 above and, for more detail, the documentation for the prescription drug file (HC-229A) when analyzing prescription drug spending over time.

As always, it is recommended that data users review relevant sections of the documentation for descriptions of these types of changes that might affect the interpretation of changes over time before undertaking trend analyses.

Analysts may also wish to consider using statistical techniques to smooth or stabilize analyses of trends using MEPS data such as comparing pooled time periods (e.g., 1996-1997 versus 2011-2012), working with moving averages, or using modeling techniques with several consecutive years of MEPS data to test the fit of specified patterns over time.

Finally, statistical significance tests should be conducted to assess the likelihood that observed trends are not attributable to sampling variation. In addition, researchers should be aware of the impact of multiple comparisons on Type I error. Without making appropriate allowance for multiple comparisons, undertaking numerous statistical significance tests of trends increases the likelihood of concluding that a change has taken place when one has not.

References

- Bethel CD, Read D, Stein REK, Blumberg SJ, Wells N, Newacheck PW. Identifying Children with Special Health Care Needs: Development and Evaluation of a Short Screening Instrument. *Ambulatory Pediatrics* Volume 2, No. 1, January-February 2002, pp 38-48.
- Bird HR, Andrews H, et. al. Global Measures of Impairment for Epidemiologic and Clinical Use with Children and Adolescents. *International Journal of Methods in Psychiatric Research*, vol. 6, 1996, pp. 295-307.
- Bramlett, M.D., Dahlhamer, J.M., & Bose, J. (2021, September). [*Weighting Procedures and Bias Assessment for the 2020 National Health Interview Survey*](#). Centers for Disease Control and Prevention.
- Chowdhury, S.R., Machlin, S.R., Gwet, K.L. [*Sample Designs of the Medical Expenditure Panel Survey Household Component, 1996-2006 and 2007-2016*](#). Methodology Report #33. January 2019. Agency for Healthcare Research and Quality, Rockville, MD.
- [*Current Population Survey: 2021 Annual Social and Economic \(ASEC\) Supplement*](#). (2021). U.S. Census Bureau.
- Dahlhamer, J.M., Bramlett, M.D., Maitland, A., & Blumberg, S.J. (2021). [*Preliminary evaluation of nonresponse bias due to the COVID-19 pandemic on National Health Interview Survey estimates, April-June 2020*](#). National Center for Health Statistics.
- Ding, Y. and Hill, S.C. (2022). Evaluating Alternative Benchmarks to Improve Identification of Outlier Drug Prices for MEPS Prescribed Medicines (PMED) Data Editing. Working Paper (forthcoming). Agency for Healthcare Research and Quality, Rockville, MD.
- Fay, R.E. (1989). Theory and Application of Replicate Weighting for Variance Calculations. Proceedings of the Survey Research Methods Sections, ASA, 212-217.
- Kessler, R.C., Andrews, G., Colpe, L.J., Hiripi, E., Mroczek, D.K., Normand, S.L., Walters, E.E., and Zaslavsky, A.M. (2002). Short screening scales to monitor population prevalence and trends in non-specific psychological distress. *Psychological Medicine* 32: 959-976.
- Kroenke, K., Spitzer, R.L., and Williams, J.B. (2003). The Patient Health Questionnaire-2: Validity of a two-item depressive screener. *Medical Care* 41: 1284-1292.
- Lau, D.T., Sosa, P., Dasgupta, N., & He, H. (2021). [*Impact of the COVID-19 Pandemic on Public Health Surveillance and Survey Data Collections in the United States*](#). *American Journal of Public Health*, 111 (12), pp. 2118-2121.
- Rothbaum, J. & Bee, A. (2021, May 3). [*Coronavirus Infects Surveys, Too: Survey Nonresponse Bias and the Coronavirus Pandemic*](#). U.S. Census Bureau.

Rothbaum, J. & Bee, A. (2022, September 13). [How Has the Pandemic Continued to Affect Survey Response? Using Administrative Data to Evaluate Nonresponse in the 2022 Current Population Survey Annual Social and Economic Supplement](#). U.S. Census Bureau.

Selim A, Rogers W, Qian S, Rothendler JA, Kent EE, Kazis LE. A new algorithm to build bridges between two patient-reported health outcome instruments: the MOS SF-36® and the VR-12 Health Survey. *Qual Life Res.* 2018 Aug; 27(8):2195-2206. View Related Profiles.

Selim AJ, Rogers W, Fleishman JA, Qian SX, Fincke BG, Rothendler JA, Kazis LE. Updated U.S. population standard for the Veterans RAND 12-item Health Survey (VR-12). *Qual Life Res.* 2009 Feb; 18(1):43-52.

Zuvekas, S.H. & Kashihara, D. (2021). [The Impacts of the COVID-19 Pandemic on the Medical Expenditure Panel Survey](#). *American Journal of Public Health, 111 (12)*, pp. 2157-2166.

D. Variable-Source Crosswalk

FOR MEPS HC-233: 2021 CONSOLIDATED DATA FILE

SURVEY ADMINISTRATION VARIABLES - PUBLIC USE

VARIABLE	DESCRIPTION	SOURCE
DUID	Panel # + Encrypted DU Identifier	Assigned in Sampling
PID	Person Number	Assigned in Sampling or by CAPI
DUPERSID	Person ID (DUID + PID)	Assigned in Sampling
PANEL	Panel Number	Constructed
FAMID31	Family ID (Student Merged In) - R3/1	CAPI Derived
FAMID42	Family ID (Student Merged In) - R4/2	CAPI Derived
FAMID53	Family ID (Student Merged In) - R5/3	CAPI Derived
FAMID21	Family ID (Student Merged In) - 12/31/21	CAPI Derived
FAMIDYR	Annual Family Identifier	Constructed
CPSFAMID	CPS-Like Family Identifier	Constructed
FCSZ1231	Family Size Responding 12/31 CPS Family	Constructed
FCRP1231	Ref Person of 12/31 CPS Family	Constructed
RULETR31	RU Letter - R3/1	CAPI Derived
RULETR42	RU Letter - R4/2	CAPI Derived
RULETR53	RU Letter - R5/3	CAPI Derived
RULETR21	RU Letter as of 12/31/21	CAPI Derived
RUSIZE31	RU Size - R3/1	CAPI Derived
RUSIZE42	RU Size - R4/2	CAPI Derived
RUSIZE53	RU Size - R5/3	CAPI Derived
RUSIZE21	RU Size as of 12/31/21	CAPI Derived
RUCLAS31	RU fielded as: Standard/New/Student - R3/1	CAPI Derived
RUCLAS42	RU fielded as: Standard/New/Student - R4/2	CAPI Derived
RUCLAS53	RU fielded as: Standard/New/Student - R5/3	CAPI Derived
RUCLAS21	RU fielded as: Standard/New/Student-12/31/21	CAPI Derived
FAMSZE31	RU Size Including Students - R3/1	CAPI Derived
FAMSZE42	RU Size Including Students - R4/2	CAPI Derived

VARIABLE	DESCRIPTION	SOURCE
FAMSZE53	RU Size Including Students - R5/3	CAPI Derived
FAMSZE21	RU Size Including Students as of 12/31/21	CAPI Derived
FMRS1231	Member of Responding 12/31 Family	Constructed
FAMS1231	Family Size of Responding 12/31 Family	Constructed
FAMSZEYR	Size of Responding Annualized Family	Constructed
FAMRFPYR	Reference Person of Annualized Family	Constructed
REGION31	Census Region - R3/1	Assigned in Sampling
REGION42	Census Region - R4/2	Assigned in Sampling
REGION53	Census Region - R5/3	Assigned in Sampling
REGION21	Census Region as of 12/31/21	Assigned in Sampling
REFPRS31	Reference Person at - R3/1	RE480-RE500
REFPRS42	Reference Person at - R4/2	RE480-RE500
REFPRS53	Reference Person at - R5/3	RE480-RE500
REFPRS21	Reference Person as of 12/31/21	RE480-RE500
RESP31	1st Respondent Indicator for R3/1	ST30
RESP42	1st Respondent Indicator for R4/2	ST30
RESP53	1st Respondent Indicator for R5/3	ST30
RESP21	1st Respondent Indicator as of 12/31/21	ST30
PROXY31	Was Respondent a Proxy in R3/1	ST30
PROXY42	Was Respondent a Proxy in R4/2	ST30
PROXY53	Was Respondent a Proxy in R5/3	ST30
PROXY21	Was Respondent a Proxy as of 12/31/21	ST30
INTVLANG	Language Interview Was Completed	RU30
BEGRFM31	R3/1 Reference Period Begin Date: Month	CAPI Derived
BEGRFY31	R3/1 Reference Period Begin Date: Year	CAPI Derived
ENDRFM31	R3/1 Reference Period End Date: Month	CAPI Derived
ENDRFY31	R3/1 Reference Period End Date: Year	CAPI Derived
BEGRFM42	R4/2 Reference Period Begin Date: Month	CAPI Derived
BEGRFY42	R4/2 Reference Period Begin Date: Year	CAPI Derived
ENDRFM42	R4/2 Reference Period End Date: Month	CAPI Derived
ENDRFY42	R4/2 Reference Period End Date: Year	CAPI Derived
BEGRFM53	R5/3 Reference Period Begin Date: Month	CAPI Derived
BEGRFY53	R5/3 Reference Period Begin Date: Year	CAPI Derived

VARIABLE	DESCRIPTION	SOURCE
ENDRFM53	R5/3 Reference Period End Date: Month	CAPI Derived
ENDRFY53	R5/3 Reference Period End Date: Year	CAPI Derived
ENDRFM21	2021 Reference Period End Date: Month	RE Section
ENDRFY21	2021 Reference Period End Date: Year	RE Section
KEYNESS	Person Key Status	RE Section
INSCOP31	Inscope - R3/1	RE Section
INSCOP42	Inscope - R4/2	RE Section
INSCOP53	Inscope - R5/3	RE Section
INSCOP21	Inscope - R5/3 Start through 12/31/21	RE Section
INSC1231	Inscope Status on 12/31/21	Constructed
INSCOPE	Was Person Ever Inscope in 2021	RE Section
ELGRND31	Eligibility - R3/1	RE Section
ELGRND42	Eligibility - R4/2	RE Section
ELGRND53	Eligibility - R5/3	RE Section
ELGRND21	Eligibility Status as of 12/31/21	RE Section
PSTATS31	Person Disposition Status - R3/1	RE Section
PSTATS42	Person Disposition Status - R4/2	RE Section
PSTATS53	Person Disposition Status - R5/3	RE Section
RURSLT31	RU Result - R3/1	Assigned by CAPI
RURSLT42	RU Result - R4/2	Assigned by CAPI
RURSLT53	RU Result - R5/3	Assigned by CAPI

DEMOGRAPHIC VARIABLES - PUBLIC USE

VARIABLE	DESCRIPTION	SOURCE
AGE31X	Age - R3/1 (Edited/Imputed)	RE80, 650, 680, 750
AGE42X	Age - R4/2 (Edited/Imputed)	RE80, 650, 680, 750
AGE53X	Age - R5/3 (Edited/Imputed)	RE80, 650, 680, 750
AGE21X	Age as of 12/31/21 (Edited/Imputed)	RE80, 650, 680, 750
AGELAST	Person's Age Last Time Eligible	AGE21X, AGE42X, AGE31X
DOBMM	Date of Birth: Month	RE70, 650, 680, 730
DOBY	Date of Birth: Year	RE70, 650, 680, 730
SEX	Sex	RE60, 650, 680, RE 720
RACEV1X	Race (Edited/Imputed)	RE1170
RACEV2X	Race (Edited/Imputed)	RE1170
RACEAX	Asian Among Races Rptd (Edited/Imputed)	RE1170
RACEBX	Black Among Races Rptd (Edited/Imputed)	RE1170
RACEWX	White Among Races Rptd (Edited/Imputed)	RE1170
RACETHX	Race/Ethnicity (Edited/Imputed)	RE1170
HISPANX	Hispanic Ethnicity (Edited/Imputed)	RE1170
HISPNCAT	Hispanic Ethnicity (Edited/Imputed)	RE1170
MARRY31X	Marital Status - R3/1 (Edited/Imputed)	RE100, 1170
MARRY42X	Marital Status - R4/2 (Edited/Imputed)	RE100, 1170
MARRY53X	Marital Status - R5/3 (Edited/Imputed)	RE100, 1170
MARRY21X	Marital Status-12/31/21 (Edited/Imputed)	RE100, 1170
SPOUID31	Spouse ID - R3/1	RE900
SPOUID42	Spouse ID - R4/2	RE900
SPOUID53	Spouse ID - R5/3	RE900
SPOUID21	Spouse ID - 12/31/21	RE900
SPOUIN31	Marital Status w/ Spouse Present - R3/1	RE900
SPOUIN42	Marital Status w/ Spouse Present - R4/2	RE900
SPOUIN53	Marital Status w/ Spouse Present - R5/3	RE900
SPOUIN21	Marital Status w/Spouse Present-12/31/21	RE900
EDUCYR	Years of Educ When First Entered MEPS	RE1180-1200
HIDEG	Highest Degree When First Entered MEPS	RE1180-1200
FTSTU31X	Student Status if Ages 17-23 - R3/1	RE50, 1210

VARIABLE	DESCRIPTION	SOURCE
FTSTU42X	Student Status if Ages 17-23 - R4/2	RE50, 1210
FTSTU53X	Student Status if Ages 17-23 - R5/3	RE50, 1210
FTSTU21X	Student Status if Ages 17-23 - 12/31/21	RE50, 1210
ACTDTY31	Military Full-Time Active Duty - R3/1	RE110, 1050, 1080, 1100
ACTDTY42	Military Full-Time Active Duty - R4/2	RE110, 1050, 1080, 1100
ACTDTY53	Military Full-Time Active Duty - R5/3	RE110, 1050, 1080, 1100
REFRL31X	Relation to Ref Pers - R3/1 (Edit/Imp)	RE 900
REFRL42X	Relation to Ref Pers - R4/2 (Edit/Imp)	RE 900
REFRL53X	Relation to Ref Pers - R5/3 (Edit/Imp)	RE 900
REFRL21X	Relation to Ref Pers - 12/31/21 (Edit/Imp)	RE 900
OTHLGSPK	Speak Other Language at Home	RE 1170
WHTLGSPK	What Language Spoken Other Than English	RE 1170
HWELLSPK	How Well Person Speaks English	RE 1170
BORNUSA	Person Born in the US	RE1170
YRSINUS	Years Person Lived in the US	RE1170
MOPID31X	PID of Person's Mom - RD 3/1	RE 900
MOPID42X	PID of Person's Mom - RD 4/2	RE 900
MOPID53X	PID of Person's Mom - RD 5/3	RE 900
DAPID31X	PID of Person's Dad - RD 3/1	RE 900
DAPID42X	PID of Person's Dad - RD 4/2	RE 900
DAPID53X	PID of Person's Dad - RD 5/3	RE 900

INCOME VARIABLES

VARIABLE	DESCRIPTION	SOURCE
FILEDR21	Has Person Filed A Fed Income Tax Return	IN 20
WILFIL21	Will Person File Fed Income Tax Return	IN 30
FLSTAT21	Person's Filing Status	IN 40
FILER21	Primary Or Secondary Filer	IN 40
JTINRU21	Joint Filer's Membership In RU	IN 50
JNTPID21	PID of Joint Filer	IN 50
TAXFRM21	Tax Form Person Will File	IN 60
FOODST21	Did Anyone Receive Food Stamps	IN 700
FOODMN21	Number Of Months Food Stamps Received	IN 710
FOODVL21	Monthly Value Of Food Stamps	IN 720
TTLP21X	Person's Total Income	Constructed
FAMINC21	Family's Total Income	Constructed
POVCAT21	Family Income As Percent Of Poverty Line - Categorical	Constructed
POVLEV21	Family Income As Percent Of Poverty Line - Continuous	Constructed
WAGEP21X	Person's Wage Income	Constructed
WAGIMP21	Wage Imputation Flag	Constructed
BUSNP21X	Person's Business Income	Constructed
BUSIMP21	Business Income Imputation Flag	Constructed
UNEMP21X	Person's Unemployment Comp Income	Constructed
UNEIMP21	Unemployment Imputation Flag	Constructed
WCMPP21X	Person's Workers' Compensation	Constructed
WCPIMP21	Workers' Comp Imputation Flag	Constructed
INTRP21X	Person's Interest Income	Constructed
INTIMP21	Interest Imputation Flag	Constructed
DIVDP21X	Person's Dividend Income	Constructed
DIVIMP21	Dividend Imputation Flag	Constructed
SALEP21X	Person's Sales Income	Constructed
SALIMP21	Sales Income Imputation Flag	Constructed
PENSP21X	Person's Pension Income	Constructed
PENIMP21	Pension Income Imputation Flag	Constructed
SSECP21X	Person's Social Security Income	Constructed

VARIABLE	DESCRIPTION	SOURCE
SSCIMP21	Social Security Imputation Flag	Constructed
TRSTP21X	Person's Trust/Rent Income	Constructed
TRTIMP21	Trust Income Imputation Flag	Constructed
VETSP21X	Person's Veteran's Income	Constructed
VETIMP21	Veteran's Income Imputation Flag	Constructed
IRASP21X	Person's Ira Income	Constructed
IRAIMP21	Ira Income Imputation Flag	Constructed
ALIMP21X	Person's Alimony Income	Constructed
ALIIMP21	Alimony Income Imputation Flag	Constructed
CHLDP21X	Person's Child Support	Constructed
CHLIMP21	Child Support Imputation Flag	Constructed
CASHP21X	Person's Other Regular Cash Contrib	Constructed
CSHIMP21	Cash Contribution Imputation Flag	Constructed
SSIP21X	Person's SSI	Constructed
SSIIMP21	SSI Imputation Flag	Constructed
PUBP21X	Person's Public Assistance	Constructed
PUBIMP21	Public Assistance Imputation Flag	Constructed
OTHRP21X	Person's Other Income	Constructed
OTHIMP21	Other Income Imputation Flag	Constructed
HIEUIDX	Health Insurance Eligibility Unit Identifier	Constructed

PERSON-LEVEL CONDITION VARIABLES - PUBLIC USE

VARIABLE	DESCRIPTION	SOURCE
RTHLTH31	Perceived Health Status - RD 3/1	PE10
RTHLTH42	Perceived Health Status - RD 4/2	PE10
RTHLTH53	Perceived Health Status - RD 5/3	PE10
MNHLTH31	Perceived Mental Health Status - RD 3/1	PE20
MNHLTH42	Perceived Mental Health Status - RD 4/2	PE20
MNHLTH53	Perceived Mental Health Status - RD 5/3	PE20
HIBPDX	High Blood Pressure Diag (>17)	PE30A
HIBPAGED	Age of Diagnosis-High Blood Pressure	PE40
BPMLDX	Mult Diag High Blood Press (>17)	PE50
CHDDX	Coronary Hrt Disease Diag (>17)	PE30B
CHDAGED	Age of Diagnosis-Coronary Heart Disease	PE60
ANGIDX	Angina Diagnosis (>17)	PE30C
ANGIAGED	Age of Diagnosis-Angina	PE70
MIDX	Heart Attack (MI) Diag (>17)	PE30D
MIAGED	Age of Diagnosis-Heart Attack (MI)	PE80
OHRTDX	Other Heart Disease Diag (>17)	PE30E
OHRTAGED	Age of Diagnosis-Other Heart Disease	PE100
OHRRTYPE	Type of Other Heart Disease (>17)	PE90
STRKDX	Stroke Diagnosis (>17)	PE30F
STRKAGED	Age of Diagnosis-Stroke	PE110
EMPHDX	Emphysema Diagnosis (>17)	PE30G
EMPHAGED	Age of Diagnosis-Emphysema	PE120
CHBRON31	Chronc Bronchits Last 12 Mths (>17)-R3/1	PE330
CHBRON53	Chronc Bronchits Last 12 Mths (>17)-R5/3	PE330
CHOLDX	High Cholesterol Diagnosis (>17)	PE30H
CHOLAGED	Age of Diagnosis-High Cholesterol	PE130
CANCERDX	Cancer Diagnosis (>17)	PE30I
CABLADDR	Cancer Diagnosed - Bladder (>17)	PE140
CABREAST	Cancer Diagnosed - Breast (>17)	PE140
CACERVIX	Cancer Diagnosed - Cervical (>17)	PE140
CACOLON	Cancer Diagnosed - Colon (>17)	PE140

VARIABLE	DESCRIPTION	SOURCE
CALUNG	Cancer Diagnosed - Lung (>17)	PE140
CALYMPH	Cancer Diagnosed - Lymphoma (Non-Hodgkin's) (>17)	PE140
CAMELANO	Cancer Diagnosed - Skin Melanoma (>17)	PE140
CAOTHER	Cancer Diagnosed - Other (>17)	PE140
CAPROSTA	Cancer Diagnosed - Prostate (>17)	PE140
CASKINNM	Cancer Diagnosed - Skin-Nonmelano (>17)	PE140
CASKINDK	Cancer Diagnosed - Skin-Unknown Type (>17)	PE140
CAUTERUS	Cancer Diagnosed - Uterine (>17)	PE140
DIABDX_M18	Diabetes Diagnosis	PE30K
DIABAGED	Age of Diagnosis-Diabetes	PE190
JTPAIN31_M18	Joint Pain Last 12 Months (>17) - RD 3/1	PE320
JTPAIN53_M18	Joint Pain Last 12 Months (>17) - RD 5/3	PE320
ARTHDX	Arthritis Diagnosis (>17)	PE30J
ARTHTYPE	Type Of Arthritis Diagnosed (>17)	PE170
ARTHAGED	Age of Diagnosis-Arthritis	PE180
ASTHDX	Asthma Diagnosis	PE30L
ASTHAGED	Age of Diagnosis-Asthma	PE200
ASSTIL31	Does Person Still Have Asthma - RD 3/1	PE210
ASSTIL53	Does Person Still Have Asthma - RD 5/3	PE210
ASATAK31	Asthma Attack Last 12 Mos- RD 3/1	PE220
ASATAK53	Asthma Attack Last 12 Mos- RD 5/3	PE220
ASTHEP31	When Was Last Episode Of Asthma - RD 3/1	PE230
ASTHEP53	When Was Last Episode Of Asthma - RD 5/3	PE230
ASACUT31	Used Acute Pres Inhaler Last 3 Mos - RD 3/1	PE260
ASACUT53	Used Acute Pres Inhaler Last 3 Mos - RD 5/3	PE260
ASMRCN31	Used >3 Acute Cn Pres Inh Last 3 Mos - RD 3/1	PE270
ASMRCN53	Used >3 Acute Cn Pres Inh Last 3 Mos - RD 5/3	PE270
ASPREV31	Ever Used Prev Daily Asthma Meds - RD 3/1	PE240
ASPREV53	Ever Used Prev Daily Asthma Meds - RD 5/3	PE240
ASDALY31	Now Take Prev Daily Asthma Meds - RD 3/1	PE250
ASDALY53	Now Take Prev Daily Asthma Meds - RD 5/3	PE250
ASPKFL31	Have Peak Flow Meter At Home - RD 3/1	PE280
ASPKFL53	Have Peak Flow Meter At Home - RD 5/3	PE280

VARIABLE	DESCRIPTION	SOURCE
ASEVFL31	Ever Used Peak Flow Meter - RD 3/1	PE290
ASEVFL53	Ever Used Peak Flow Meter - RD 5/3	PE290
ASWNFL31	When Last Used Peak Flow Meter - RD 3/1	PE300
ASWNFL53	When Last Used Peak Flow Meter - RD 5/3	PE300
ADHDADDX	ADHD/ADD Diagnosis (5-17)	PE30M
ADHDAGED	Age of Diagnosis-ADHD/ADD	PE310

HEALTH STATUS VARIABLES - PUBLIC USE

VARIABLE	DESCRIPTION	SOURCE
IADLHP31	IADL Screener - RD 3/1	HE10-30
IADLHP53	IADL Screener - RD 5/3	HE10-30
ADLHLP31	ADL Screener - RD 3/1	HE40-60
ADLHLP53	ADL Screener - RD 5/3	HE40-60
AIDHLP31	Used Assistive Devices - RD 3/1	HE70-80
AIDHLP53	Used Assistive Devices - RD 5/3	HE70-80
WLKLIM31	Limitation in Physical Functioning - RD 3/1	HE90-100
WLKLIM53	Limitation in Physical Functioning - RD 5/3	HE90-100
LFTDIF31	Difficulty Lifting 10 Pounds - RD 3/1	HE110
LFTDIF53	Difficulty Lifting 10 Pounds - RD 5/3	HE110
STPDIF31	Difficulty Walking up 10 Steps - RD 3/1	HE120
STPDIF53	Difficulty Walking up 10 Steps - RD 5/3	HE120
WLKDIF31	Difficulty Walking 3 Blocks - RD 3/1	HE130
WLKDIF53	Difficulty Walking 3 Blocks - RD 5/3	HE130
MILDIF31	Difficulty Walking a Mile - RD 3/1	HE140
MILDIF53	Difficulty Walking a Mile - RD 5/3	HE140
STNDIF31	Difficulty Standing 20 Minutes - RD 3/1	HE150
STNDIF53	Difficulty Standing 20 Minutes - RD 5/3	HE150
BENDIF31	Difficulty Bending/Stooping - RD 3/1	HE160
BENDIF53	Difficulty Bending/Stooping - RD 5/3	HE160
RCHDIF31	Difficulty Reaching Overhead - RD 3/1	HE170
RCHDIF53	Difficulty Reaching Overhead - RD 5/3	HE170
FNGRDF31	Difficulty Using Fingers to Grasp - RD 3/1	HE180
FNGRDF53	Difficulty Using Fingers to Grasp - RD 5/3	HE180
ACTLIM31	Any Limitation Work/Housewrk/Schl - RD 3/1	HE190-200
ACTLIM53	Any Limitation Work/Housewrk/Schl - RD 5/3	HE190-200
WRKLIM31	Work Limitation - RD 3/1	HE210
WRKLIM53	Work Limitation - RD 5/3	HE210
HSELIM31	Housework Limitation - RD 3/1	HE210
HSELIM53	Housework Limitation - RD 5/3	HE210
SCHLIM31	School Limitation - RD 3/1	HE210

VARIABLE	DESCRIPTION	SOURCE
SCHLIM53	School Limitation - RD 5/3	HE210
UNABLE31	Completely Unable to Do Activity - RD 3/1	HE220
UNABLE53	Completely Unable to Do Activity - RD 5/3	HE220
SOCLIM31	Social Limitations - RD 3/1	HE230-240
SOCLIM53	Social Limitations - RD 5/3	HE230-240
COGLIM31	Cognitive Limitations - RD 3/1	HE250A, 250B, 250C, 260
COGLIM53	Cognitive Limitations - RD 5/3	HE250A, 250B, 250C, 260
DFHEAR42	Serious Difficulty Hearing-RD4/2	HE270-280
DFSEE42	Serious Difficulty See w/Glasses-RD4/2	HE290C-300
DFCOG42	Serious Cognitive Difficulties-RD4/2	HE310-320
DFWLKC42	Serious Difculty Wlk/Climb Stairs-RD4/2	HE330-340
DFDRSB42	Difficulty Dressing/Bathing-RD4/2	HE350-360
DFERND42	Difficulty Doing Errands Alone-RD4/2	HE370-380
ANYLMI21	Any Limitation in P23R7,8/P24R5,6,7/P25R3,4/P26R1,2,3	Constructed
CHPMED42	CSHCN: Child Needs Prescrb Med(0-17)-R4/2	CS10
CHPMHB42	CSHCN: Pmed for Hlth/Behv Cond (0-17)- R4/2	CS20
CHPMC42	CSHCN: Pmed Cond Last 12+ Mos (0-17)- R4/2	CS30
CHSERV42	CSHCN: Chld Needs Med&Oth Serv (0-17)- R4/2	CS40
CHSRHB42	CSHCN: Serv for Hlth/Behv Cond(0-17)-R4/2	CS50
CHSRC42	CSHCN: Serv Cond Last 12+ Mos (0-17)-R4/2	CS60
CHLIMI42	CSHCN: Limited in Any Way (0-17)-R4/2	CS70
CHLIHB42	CSHCN: Limt for Hlth/Behv Cond(0-17)-R4/2	CS80
CHLICO42	CSHCN: Limit Cond Last 12+ Mos (0-17)- R4/2	CS90
CHTHER42	CSHCN: Chld Needs Spec Therapy (0-17)- R4/2	CS100
CHTHHB42	CSHCN: Spec Ther for Hlth+Cond(0-17)-R4/2	CS110
CHTHCO42	CSHCN: Ther Cond Last 12+ Mos (0-17)-R4/2	CS120
CHCOUN42	CSHCN: Child Needs Counseling (0-17)-R4/2	CS130

VARIABLE	DESCRIPTION	SOURCE
CHEMPB42	CSHCN: Couns Prob Last 12+ Mos (0-17)-R4/2	CS140
CSHCN42	CSHCN:Child w/Spec HC Needs (0-17)-R4/2	CS10-140
GETTRB42	Prblm Getting in Trouble (5-17)-R4/2	CS150A
MOMPRO42	Problem Getting Along w/Mom (5-17)-R4/2	CS150B
DADPRO42	Problem Getting Along w/Dad (5-17)-R4/2	CS150C
UNHAP42	Problem Feeling Unhappy/Sad (5-17)-R4/2	CS150D
SCHLBH42	Problem Behavior at School (5-17)-R4/2	CS160A
HAVFUN42	Problem Having Fun (5-17) - R4/2	CS160B
ADUPRO42	Prblm Getting Along w/Adults (5-17)-R4/2	CS160C
NERVAF42	Prblm Feeling Nervous/Afraid (5-17)-R4/2	CS170A
SIBPRO42	Prblm Getting Along w/Sibs (5-17)-R4/2	CS170B
KIDPRO42	Prblm Getting Along w/Kids (5-17)-R4/2	CS170C
SPRPRO42	Problem w/Sports/Hobbies (5-17)-R4/2	CS180A
SCHPRO42	Problem With Schoolwork (5-17)-R4/2	CS180B
HOMEBH42	Problem w/Behavior at Home (5-17)-R4/2	CS180C
CHILCR42	CAHPS:12Mos: Ill/Inj Need Care (0-17)R4/2	CS190
CHILWW42	CAHPS:12Mos: Ill Care Whn Wntd (0-17)R4/2	CS200
CHRTCR42	CAHPS:12Mos: Make Rout Apt (0-17)R4/2	CS210
CHRTWW42	CAHPS:12Mos: Rout Apt Whn Wntd (0-17)R4/2	CS220
CHAPPT42	CAHPS:12Mos: # of Off/Clin Apts (0-17)R4/2	CS230
CHLIST42	CAHPS:12Mos: Chld Dr Lsn to You (0-17)R4/2	CS250
CHEXPL42	CAHPS:12Mos: Chld Dr Expl Thng (0-17)R4/2	CS240
CHRESP42	CAHPS:12Mos: Chld's Dr Shw Resp(0-17)R4/2	CS260
CHPRTM42	CAHPS:12Mos: Child Dr Engh Time(0-17)R4/2	CS270
CHHECR42	CAHPS:12Mos: Rate Chld Hlt Care (0-17)R4/2	CS280
CHSPEC42_M18	CAHPS:12Mos: Make Appt for Spec (0-17)R4/2	CS290
CHEYRE42_M18	CAHPS:12Mos: Get Appt for Spec (0-17)R4/2	CS300

VARIABLE	DESCRIPTION	SOURCE
LSTETH53	Lost All Uppr And Lowr Teeth (>17) - RD 5/3	AH100
PHYEXE53	Mod/Vig Phys Exec 5X Wk (>17) - RD 5/3	AH110
OFTSMK53	How Oftn Smoke Cigarettes (>17) - RD 5/3	AH120
SAQELIG	Eligibility Status for SAQ	Constructed
ADPROX42	SAQ: Relationship of Proxy to Adult	Constructed
ADGENH42	SAQ: Health in General VR-12	SAQ Q1
ADDAYA42	SAQ: Hlth Limits Mod Activities VR-12	SAQ Q2a
ADCLIM42	SAQ: Hlth Limits Climbing Stairs VR-12	SAQ Q2b
ADACLS42	SAQ 4Wks:Accmp Less B/C Phy Prbs VR-12	SAQ Q3a
ADWKLM42	SAQ 4Wks:Work Limt B/C Phy Probs VR-12	SAQ Q3b
ADEMLS42	SAQ 4Wks:Accmp Less B/C Mnt Prbs VR-12	SAQ Q4a
ADMWCF42	SAQ 4Wks:Work Limt B/C Mnt Probs VR-12	SAQ Q4b
ADPAIN42	SAQ 4Wks:Pain Limits Normal Work VR-12	SAQ Q5
ADPCFL42	SAQ 4Wks: Felt Calm/Peaceful VR-12	SAQ Q6a
ADENGY42	SAQ 4Wks: Had a Lot of Energy VR-12	SAQ Q6b
ADPRST42	SAQ 4Wks: Felt Downhearted/Depr VR-12	SAQ Q6c
ADSOCA42	SAQ 4Wks: Hlth Stopped Soc Activ VR-12	SAQ Q7
ADNERV42	SAQ 30 Days: How Often Felt Nervous	SAQ Q8a
ADHOPE42	SAQ 30 Days: How Often Felt Hopeless	SAQ Q8b
ADREST42	SAQ 30 Days: How Often Felt Restless	SAQ Q8c
ADSAD42	SAQ 30 Days: How Often Felt Sad	SAQ Q8d
ADEFRT42	SAQ 30 Days: How Oftn Everything an Effort	SAQ Q8e
ADWRTH42	SAQ 30 Days: How Often Felt Worthless	SAQ Q8f
K6SUM42	SAQ 30 Days: Overall Rating of Feelings	Constructed
ADINTR42	SAQ 2 Wks: Little Interest in Things	SAQ Q9a
ADDPRS42	SAQ 2 Wks: Felt Down/Depressed/Hopeless	SAQ Q9b
PHQ242	SAQ 2 Wks: Overall Rating of Feelings	Constructed
ADINSA42	SAQ: Do Not Need Health Insurance	SAQ Q10a
ADINSB42	SAQ: Health Insurance Not Worth Cost	SAQ Q10b
ADRISK42	SAQ: More Likely to Take Risks	SAQ Q10c
ADOVER42	SAQ: Can Overcome Ills Without Med Help	SAQ Q10d
ADILCR42	SAQ 12 Mos: Ill/Injury Needing Immed Care	SAQ Q11
ADILWW42	SAQ 12 Mos: Got Care When Needed Ill/Inj	SAQ Q12

VARIABLE	DESCRIPTION	SOURCE
ADRTCR42	SAQ 12 Mos: Made Appt Routine Med Care	SAQ Q13
ADRTWW42	SAQ 12 Mos: Got Med Appt When Needed	SAQ Q14
ADAPPT42	SAQ 12 Mos: # Visits to Med Off for Care	SAQ Q15
ADHECR42	SAQ 12 Mos: Rating of Health Care	SAQ Q16
ADINST42	SAQ 12 Mos: Dr Gave Spcific Instrctns	SAQ Q17
ADEZUN42	SAQ 12 Mos: Dr Given Instr. EZ Understd	SAQ Q18
ADTLHW42	SAQ 12 Mos: Dr Asked R Desc How Follow	SAQ Q19
ADFFRM42	SAQ 12 Mos: Had to Fill Out/Sign Forms	SAQ Q20
ADFHLP42	SAQ 12 Mos: Offrd Help Filling Out Forms	SAQ Q21
ADEXPL42	SAQ 12 Mos: Doc Explained So Understood	SAQ Q22
ADLIST42	SAQ 12 Mos: Doctor Listened to You	SAQ Q23
ADRESP42	SAQ 12 Mos: Doctor Showed Respect	SAQ Q24
ADPRTM42	SAQ 12 Mos: Dr Spent Enuf Time With You	SAQ Q25
ADSMOK42	SAQ: Currently Smoking	SAQ Q26
ADNSMK42	SAQ 12 Mos: Dr Advised to Quit Smoking	SAQ Q27
ADSPCL42	SAQ 12 Mos: Needed to See Specialist	SAQ Q28
ADSNP42	SAQ 12 Mos: How Eesy to See Specialist	SAQ Q29
VPCS42	SAQ: Phy Component Summry VR-12 Imputed	Constructed
VMCS42	SAQ: Mnt Component Summry VR-12 Imputed	Constructed
VRFLAG42	SAQ: VPCS42/VMCS42 Imputation Flag VR-12	Constructed
ADCMPM42	SAQ: Date Completed - Month	SAQ Completed Month field
ADCMPY42	SAQ: Date Completed - Year	SAQ Completed Year field
ADLANG42	SAQ: Language of SAQ Interview	Constructed
DCSELIG	DCS: Eligibility Status for DCS	Constructed
DSDIA53	DCS: Diabetes Diagnosis By Health Prof	DCS Q1
DSA1C53	DCS: Times Tested for A-One-C in 2021	DCS Q2
DSFT2253	DCS: Had Feet Checked During 2022	DCS Q3
DSFT2153	DCS: Had Feet Checked During 2021	DCS Q3
DSFT2053	DCS: Had Feet Checked During 2020	DCS Q3
DSFB2053	DCS: Had Feet Checked Before 2020	DCS Q3

VARIABLE	DESCRIPTION	SOURCE
DSFTNV53	DCS: Never Had Feet Checked	DCS Q3
DSEY2253	DCS: Dilated Eye Exam in 2022	DCS Q4
DSEY2153	DCS: Dilated Eye Exam in 2021	DCS Q4
DSEY2053	DCS: Dilated Eye Exam in 2020	DCS Q4
DSEB2053	DCS: Dilated Eye Exam Before 2020	DCS Q4
DSEYNV53	DCS: Never Had Dilated Eye Exam	DCS Q4
DSCH2253	DCS: Blood Cholesterol Checked in 2022	DCS Q5
DSCH2153	DCS: Blood Cholesterol Checked in 2021	DCS Q5
DSCH2053	DCS: Blood Cholesterol Checked in 2020	DCS Q5
DSCB2053	DCS: Blood Cholesterol Checked Before 2020	DCS Q5
DSCHNV53	DCS: Never Had Blood Cholesterol Checked	DCS Q5
DSFL2253	DCS: Got Flu Vaccination in 2022	DCS Q6
DSFL2153	DCS: Got Flu Vaccination in 2021	DCS Q6
DSFL2053	DCS: Got Flu Vaccination in 2020	DCS Q6
DSVB2053	DCS: Got Flu Vaccination Before 2020	DCS Q6
DSFLNV53	DCS: Never Got Flu Vaccination	DCS Q6
DSKIDN53	DCS: Has Diabetes Caused Kidney Problems	DCS Q7
DSEYPR53	DCS: Has Diabetes Caused Eye Probs	DCS Q6
DSDIET53	DCS: Treat Diabetes w/Diet Modification	DCS Q9
DSMED53	DCS: Treat Diabetes w/Meds by Mouth	DCS Q10
DSINSU53	DCS: Treat Diabetes w/Insulin Injections	DCS Q11
DSCPCP53	DCS: Learned Diab Care from Prim Care Prov	DCS Q13
DSCNPC53	DCS: Learned Diab Care from Other Prov	DCS Q13
DSCPHN53	DCS: Learned Diab Care from Phn Call w/Prov	DCS Q13
DSCINT53	DCS: Learned Diab Care from Reading Internet	DCS Q13
DSCGRP53	DCS: Learned Diab Care by Taking Grp Class	DCS Q13
DSCONF53	DSC: Confident Taking Care of Diabetes	DCS Q14
DSPRX53	DCS: Was Respondent a Proxy	Constructed
SDOHELIG	Eligibility Status for SDOH	Constructed
SDLIFE	SDOH: Satisfied with Life	SDOH Q1
SDHOME	SDOH: Satisfied with Home	SDOH Q2
SDMEDCARE	SDOH: Places for Medical Care	SDOH Q3a
SDPARKS	SDOH: Places for Parks/Play	SDOH Q3b

VARIABLE	DESCRIPTION	SOURCE
SDHLTHFOOD	SDOH: Places for Healthy Food	SDOH Q3c
SDSFCRIME	SDOH: Safe From Crime/Violence	SDOH Q3d
SDPUBTRANS	SDOH: Access Public Transportation	SDOH Q3e
SDAFRDHOME	SDOH: Affordable Housing	SDOH Q3f
SDNOTRANS	SDOH: 12 Mos No Trans Daily	SDOH Q4
SDLATERENT	SDOH: 12 Mos Pay Rent/Mortg Late	SDOH Q5
SDLATEUTIL	SDOH: 12 Mos Pay Utility Late	SDOH Q6
SDSHUTUTIL	SDOH: 12 Mos Threat Utility Off	SDOH Q7
SDPROBPEST	SDOH: Home Problem Pests	SDOH Q8a
SDPROBMOLD	SDOH: Home Problem Mold	SDOH Q8b
SDPROBLEAD	SDOH: Home Problem Lead	SDOH Q8c
SDPROBHEAT	SDOH: Home Problem Heat	SDOH Q8d
SDPROBCOOK	SDOH: Home Problem Cook	SDOH Q8e
SDPROBSMKDET	SDOH: Home Problem Smoke Detector	SDOH Q8f
SDPROBLEAKS	SDOH: Home Problem Water Leaks	SDOH Q8g
SDPROBNONE	SDOH: No Home Problems	SDOH Q8h
SDWRRYFD	SDOH: 12 Mos Worried About Food	SDOH Q9a
SDNOFOOD	SDOH: 12 Mos Food Ran Out	SDOH Q9b
SDPAYBASICS	SDOH: How Hard Pay Basics	SDOH Q10
SDUNEXPEXP	SDOH: Cover Unexpected Expense	SDOH Q11
SDMISSCCLN	SDOH: 12 Mos Miss Card or Loan Pymt	SDOH Q12
SDDEBT	SDOH: 12 Mos Contact Collection	SDOH Q13
SDDAYEXER	SDOH: 30 Days Avg Days Per Wk Mod Exrcs	SDOH Q14
SDMINSEXER	SDOH: 30 Days Min Per Day Mod Exrcs	SDOH Q15
SDSTRESS	SDOH: How Often Stress	SDOH Q16
SDFAMILY	SDOH: Help From Family	SDOH Q17a
SDFRIENDS	SDOH: Help From Friends	SDOH Q17b
SDCOMM	SDOH: Help From Community	SDOH Q17c
SDTLKPHN	SDOH: Telephone Others Per Week	SDOH Q18
SDGETTGT	SDOH: See Others Per Week	SDOH Q19
SDCHURCH	SDOH: Attend Church or Services	SDOH Q20
SDCLUBORG	SDOH: Meetings Per Year	SDOH Q21
SDCOMPAN	SDOH: Feel Lack Companionship	SDOH Q22a

VARIABLE	DESCRIPTION	SOURCE
SDLEFTOUT	SDOH: Feel Left Out	SDOH Q22b
SDISOL	SDOH: Feel Isolated	SDOH Q22c
SDENICPROD	SDOH: Used Electronic Nicotine Prod	SDOH Q23
SDDSCRMDR	SDOH: Discrimination in Healthcare	SDOH Q24a
SDDSCRMWRK	SDOH: Discrimination at Work	SDOH Q24b
SDDSCRMJOB	SDOH: Discrimination Applying Jobs	SDOH Q24c
SDDSCRMHS	SDOH: Discrimination in Housing	SDOH Q24d
SDDSCRMPOL	SDOH: Discrimination by Police	SDOH Q24e
SDDSCRMPUB	SDOH: Discrimination Applying Services	SDOH Q24f
SDDSCRMSTR	SDOH: Discrimination Stores	SDOH Q24g
SDPHYSHURT	SDOH: How Often Hurt by Others	SDOH Q25a
SDINSULT	SDOH: How Often Insulted	SDOH Q25b
SDTHRHARM	SDOH: How Often Threatened Harm	SDOH Q25c
SDSCREAM	SDOH: How Often Scream Curse	SDOH Q25d
SDHMDEPR	SDOH: Lived with Mental Illness <18	SDOH Q26
SDHMALC	SDOH: Lived with Alcoholic <18	SDOH Q27
SDHMDRG	SDOH: Lived with Drugs <18	SDOH Q28
SDHMJAIL	SDOH: Lived with Sentenced <18	SDOH Q29
SDHMDIV	SDOH: Lived with Split Home <18	SDOH Q30
SDHMBEAT	SDOH: How Often Abused <18	SDOH Q31
SDHURTCHLD	SDOH: How Often Child Hurt	SDOH Q32
SDINSCHLD	SDOH: How Often Child Insult	SDOH Q33
SDTCHCHLD	SDOH: How Often Child Touched	SDOH Q34
SDTCHADLT	SDOH: How Often Asked to Touch	SDOH Q35
SDFRCSXCH	SDOH: How Often Forced	SDOH Q36
SDCMPM	SDOH: Date Completed - Month	SDOH Completed Month field
SDCMPY	SDOH: Date Completed - Year	SDOH Completed Year field
SDPROX	SDOH: Relationship Respondent to Adult	SDOH Q37
SDOHELIG	Eligibility Status for SDOH	Constructed

DISABILITY DAYS VARIABLES - PUBLIC USE

VARIABLE	DESCRIPTION	SOURCE
DDNWRK21	# Days Missed Work Due to Ill/Inj 2021	AH10, 20, 30
OTHDYS21	Miss Any Work Day to Care for Oth 2021	AH40
OTHNDD21	# Day Missed Work to Care for Oth 2021	AH50, 60

ACCESS TO CARE VARIABLES - PUBLIC USE

VARIABLE	DESCRIPTION	SOURCE
ACCELI42	Pers Eligible for Access Supplement-R4/2	Constructed
HAVEUS42	Does Person Have USC Provider-R4/2	AC10, 20
PRACTP42	Provider Own/Group Practice-R4/2	AC20, 30
YNOUSC42_M18	Main Reas Pers Doesnt Have USC-R4/2	AC40
PROVTY42_M18	Provider Type - R4/2	PV30LU, 70
PLCTYP42	USC Type of Place - R4/2	AC50
TMTKUS42	How Long It Takes Get to USC-R4/2	AC60
TYPEPE42	USC Type of Provider - R4/2	AC70, 80, 90
LOCATN42	USC Location - R4/2	Constructed
HSPLAP42	Is Provider Hispanic or Latino-R4/2	AC100
WHITPR42	Is Provider White - R4/2	AC110
BLCKPR42	Is Provider Black/African Amer-R4/2	AC110
ASIANP42	Is Provider Asian - R4/2	AC110
NATAMP42	Is Provider Native American - R4/2	AC110
PACISP42	Is Provider Oth Pacific Islndr-R4/2	AC110
OTHRCP42	Is Provider Some Other Race - R4/2	AC110
GENDRP42	Is Provider Male or Female - R4/2	AC120
PHNREG42	How Diff Contact USC By Phone-R4/2	AC140
OFFHOU42	USC Has Office Hrs Nghts/Wkends-R4/2	AC150
AFTHOU42	How Diff Contact USC Aft Hours-R4/2	AC160
TREATM42	Prov Ask About Oth Treatments-R4/2	AC170
DECIDE42	Prov Asks Pers to Help Decide-R4/2	AC180
EXPLOP42	Prov Explns Options to Pers - R4/2	AC190
PRVSPK42	Prov Speaks Person's Language-R4/2	AC200
DLAYCA42	Delay Med Care For Cost-R4/2	AC210, 220
AFRDCA42	Could Not Afford Med Care-R4/2	AC230, 240
DLAYDN42	Delay Getting Dental for Cost-R4/2	AC250, 260
AFRDDN42	Could Not Afford Dental Care-R4/2	AC270, 280
DLAYPM42	Delay Getting Pmed For Cost-R4/2	AC290, 300
AFRDPM42	Could Not Afford Pmed Care-R4/2	AC310, 320
CVDLAYCA31	Delay Med Care For COVID R3/1	CV10, 20

VARIABLE	DESCRIPTION	SOURCE
CVDLAYCA42	Delay Med Care For COVID R4/2	CV10, 20
CVDLAYCA53	Delay Med Care For COVID R5/3	CV10, 20
CVDLAYDN31	Delay Med Care For COVID R3/1	CV30, 40
CVDLAYDN42	Delay Med Care For COVID R4/2	CV30, 40
CVDLAYDN53	Delay Getting Dental For COVID R5/3	CV30, 40
CVDLAYPM31	Delay Getting PMED For COVID R3/1	CV50, 60
CVDLAYPM42	Delay Getting PMED For COVID R4/2	CV50, 60
CVDLAYPM53	Delay Getting Pmed For COVID R5/3	CV50, 60
CVVACCINE42	Ever Had COVID-19 Vaccine R4/2	CV70, CV80
CVVACCINE53	Ever Had COVID-19 Vaccine R5/3	CV70, CV80
BOOSTERSHOT53	Ever Had COVID-19 Booster Shot R5/3	CV90

EMPLOYMENT VARIABLES - PUBLIC USE

VARIABLE	DESCRIPTION	SOURCE
EMPST31	Employment Status RD 3/1	EM10-30; RJ10, 60
EMPST42	Employment Status RD 4/2	EM10-30; RJ10, 60
EMPST53	Employment Status RD 5/3	EM10-30; RJ10, 60
RNDFLG31	Data Collection Round for RD 3/1 CMJ	RJ10; Constructed
MORJOB31	Has More than One Job RD 3/1 Int Date	EM40, 230; RJ10, 60; Constructed
MORJOB42	Has More than One Job RD 4/2 Int Date	EM40, 230; RJ10, 60; Constructed
MORJOB53	Has More than One Job RD 5/3 Int Date	EM40, 230; RJ10, 60; Constructed
EVRWRK	Ever Wrkd for Pay in Life as of 12/31/21	EM300; Constructed
HRWG31X	Hourly Wage RD 3/1 CMJ (Imp)	EW10, 20, 30, 40, 50, 60, 70, 110, 120, 130, 140, 180, 190, 260, 270, 280, 290, EM420, 540, 550
HRWG42X	Hourly Wage RD 4/2 CMJ (Imp)	EW10, 20, 30, 40, 50, 60, 70, 110, 120, 130, 140, 180, 190, 260, 270, 280, 290, EM420, 540, 550
HRWG53X	Hourly Wage RD 5/3 CMJ (Imp)	EW10, 20, 30, 40, 50, 60, 70, 110, 120, 130, 140, 180, 190, 260, 270, 280, 290, EM420, 540, 550
HRWGIM31	HRWG31X Imputation Flag	Constructed
HRWGIM42	HRWG42X Imputation Flag	Constructed
HRWGIM53	HRWG53X Imputation Flag	Constructed
HRHOW31	How Hourly Wage Was Calculated RD 3/1	EM420, 540, EW20-220, 260, 270, 280, 290
HRHOW42	How Hourly Wage Was Calculated RD 4/2	EM420, 540, EW20-220, 260, 270, 280, 290

VARIABLE	DESCRIPTION	SOURCE
HRHOW53	How Hourly Wage Was Calculated RD 5/3	EM420, 540, EW20-220, 260, 270, 280, 290
DIFFWG31	Persons Wages Different this RD 3/1 at CMJ	EM420, RJ10, 30
DIFFWG42	Persons Wages Different this RD 4/2 at CMJ	EM420, RJ10, 30
DIFFWG53	Persons Wages Different this RD 5/3 at CMJ	EM420, RJ10, 30
NHRWG31	Updated Hrly Wage RD 3/1 CMJ (Edited)	EW10, 20, 30, 40, 50 60, 70, 110, 120, 130, 140, 180, 190, 260, 270, 280, 290; RJ10, 30, 60, EM420, 540, 550
NHRWG42	Updated Hrly Wage RD 4/2 CMJ (Edited)	EW10, 20, 30, 40, 50 60, 70, 110, 120, 130, 140, 180, 190, 260, 270, 280, 290; RJ10, 30, 60, EM420, 540, 550
NHRWG53	Updated Hrly Wage RD 5/3 CMJ (Edited)	EW10, 20, 30, 40, 50 60, 70, 110, 120, 130, 140, 180, 190, 260, 270, 280, 290; RJ10, 30, 60, EM420, 540, 550
HOUR31	Hours Per Week at RD 3/1 CMJ	EM540-550
HOUR42	Hours Per Week at RD 4/2 CMJ	EM540-550, RJ10
HOUR53	Hours Per Week at RD 5/3 CMJ	EM540-550, RJ10
TEMPJB31	Is CMJ a Temporary Job RD 3/1	EM560
TEMPJB42	Is CMJ a Temporary Job RD 4/2	EM560
TEMPJB53	Is CMJ a Temporary Job RD 5/3	EM560
SSNLJB31	Is CMJ a Seasonal Job RD 3/1	EM570
SSNLJB42	Is CMJ a Seasonal Job RD 4/2	EM570
SSNLJB53	Is CMJ a Seasonal Job RD 5/3	EM570
SELFCM31	Self-Employed at RD 3/1 CMJ	EM420
SELFCM42	Self-Employed at RD 4/2 CMJ	EM420, RJ10
SELFCM53	Self-Employed at RD 5/3 CMJ	EM420, RJ10

VARIABLE	DESCRIPTION	SOURCE
DISVW31X	Disavowed Health Ins at RD 3/1 CMJ (Ed)	EM480, 660-670, 710-720, 740; RJ70, 80; HX and OE Sections
DISVW42X	Disavowed Health Ins at RD 4/2 CMJ (Ed)	EM480, 660-670, 710-720, 740; RJ70, 80; HX and OE Sections
DISVW53X	Disavowed Health Ins at RD 5/3 CMJ (Ed)	EM480, 660-670, 710-720, 740; RJ70, 80; HX and OE Sections
CHOIC31	Choice of Health Plans at RD 3/1 CMJ	EM480, 660 - 680, 740
CHOIC42	Choice of Health Plans at RD 4/2 CMJ	EM480, 660 - 680, 740, RJ10
CHOIC53	Choice of Health Plans at RD 5/3 CMJ	EM480, 660 - 680, 740, RJ10
INDCAT31	Industry Group RD 3/1 CMJ	EM50, 80, 110, 240, 480, 490, 500, 510; RJ10; Constructed
INDCAT42	Industry Group RD 4/2 CMJ	EM50, 80, 110, 240, 480, 490, 500, 510; RJ10; Constructed
INDCAT53	Industry Group RD 5/3 CMJ	EM50, 80, 110, 240, 480, 490, 500, 510; RJ10; Constructed
NUMEMP31	Number of Employees at RD 3/1 CMJ	EM420, 430-440, 740; Imputed
NUMEMP42	Number of Employees at RD 4/2 CMJ	EM420, 430-440, 740; Imputed
NUMEMP53	Number of Employees at RD 5/3 CMJ	EM420, 430-440, 740; Imputed
MORE31	RD 3/1 CMJ Firm Has More than 1 Location	EM420, 450
MORE42	RD 4/2 CMJ Firm Has More than 1 Location	EM10-30, 230, 420, 450; RJ10
MORE53	RD 5/3 CMJ Firm Has More than 1 Location	EM10-30, 230, 420, 450; RJ10
UNION31	Union Status at RD 3/1 CMJ	EM10-30, 230, 480, 700; RJ10

VARIABLE	DESCRIPTION	SOURCE
UNION42	Union Status at RD 4/2 CMJ	EM10-30, 230, 480, 700; RJ10
UNION53	Union Status at RD 5/3 CMJ	EM10-30, 230, 480, 700; RJ10
NWK31	Reason Not Working During RD 3/1	EM300, 750
NWK42	Reason Not Working During RD 4/2	EM300, 750
NWK53	Reason Not Working During RD 5/3	EM300, 750
CHGJ3142	Changed Job between RD 3/1 and RD 4/2	RJ10
CHGJ4253	Changed Job between RD 4/2 and RD 5/3	RJ10
YCHJ3142	Why Chngd Job between RD 3/1 and RD 4/2	RJ10, 130
YCHJ4253	Why Chngd Job between RD 4/2 and RD 5/3	RJ10, 130
STJBMM31	Month Started RD 3/1 CMJ	EM60_01-02, 90_01-02, 110_01-02, 250_01-02; RJ10
STJBYY31	Year Started RD 3/1 CMJ	EM60_01, 90_01, 110_01, 250_01; RJ10
STJBMM42	Month Started RD 4/2 CMJ	EM60_01-02, 90_01-02, 110_01-02, 250_01-02; RJ10
STJBYY42	Year Started RD 4/2 CMJ	EM60_01, 90_01, 110_01, 250_01; RJ10
STJBMM53	Month Started RD 5/3 CMJ	EM60_01-02, 90_01-02, 110_01-02, 250_01-02; RJ10
STJBYY53	Year Started RD 5/3 CMJ	EM60_01, 90_01, 110_01, 250_01; RJ10
EVRETIRE	Person Has Ever Retired	EM350, Constructed
OCCCAT31	Occupation Group RD 3/1 CMJ	EM480, 490, 500, 510; Constructed
OCCCAT42	Occupation Group RD 4/2 CMJ	EM480, 490, 500, 510; RJ10; Constructed
OCCCAT53	Occupation Group RD 5/3 CMJ	EM480, 490, 500, 510; RJ10; Constructed

VARIABLE	DESCRIPTION	SOURCE
PAYVAC31	Paid Vacation at RD 3/1 CMJ	EM420, 600
PAYVAC42	Paid Vacation at RD 4/2 CMJ	EM420, 600; RJ10
PAYVAC53	Paid Vacation at RD 5/3 CMJ	EM420, 600; RJ10
SICPAY31	Paid Sick Leave at RD 3/1 CMJ	EM420, 580
SICPAY42	Paid Sick Leave at RD 4/2 CMJ	EM420, 580; RJ10
SICPAY53	Paid Sick Leave at RD 5/3 CMJ	EM420, 580; RJ10
PAYDR31	Paid Leave to Visit Dr RD 3/1 CMJ	EM420, 580-590
PAYDR42	Paid Leave to Visit Dr RD 4/2 CMJ	EM420, 580-590; RJ10
PAYDR53	Paid Leave to Visit Dr RD 5/3 CMJ	EM420, 580-590; RJ10
RETPLN31	Pension Plan at RD 3/1 CMJ	EM420, 610
RETPLN42	Pension Plan at RD 4/2 CMJ	EM420, 610; RJ10
RETPLN53	Pension Plan at RD 5/3 CMJ	EM420, 610; RJ10
BSNTY31	Sole Prop, Partner, Corp, RD 3/1 CMJ	EM420 ,460-470
BSNTY42	Sole Prop, Partner, Corp, RD 4/2 CMJ	EM420, 460-470; RJ10
BSNTY53	Sole Prop, Partner, Corp, RD 5/3 CMJ	EM420, 460-470; RJ10
JOBORG31	Priv (Profit,Nonprofit) Gov RD 3/1 CMJ	EM420, 480
JOBORG42	Priv (Profit,Nonprofit) Gov RD 4/2 CMJ	EM420, 480; RJ10
JOBORG53	Priv (Profit,Nonprofit) Gov RD 5/3 CMJ	EM420, 480; RJ10
HELD31X	Health Insur Held from RD 3/1 CMJ (Ed)	EM480, 660, 670, 710, 740; RJ70, 80, 110; HX, HP, and OE Sections
HELD42X	Health Insur Held from RD 4/2 CMJ (Ed)	EM480, 660, 670, 710, 740; RJ70, 80, 110; HX, HP, and OE Sections
HELD53X	Health Insur Held from RD 5/3 CMJ (Ed)	EM480, 660, 670, 710, 740; RJ70, 80, 110; HX, HP, and OE Sections
OFFER31X	Health Insur Offered by RD 3/1 CMJ (Ed)	EM480, 660, 670, 710, 740; RJ70-90, 110, HX Sections

VARIABLE	DESCRIPTION	SOURCE
OFFER42X	Health Insur Offered by RD 4/2 CMJ (Ed)	EM480, 660, 670, 710, 740; RJ70-90, 110, HX Sections
OFFER53X	Health Insur Offered by RD 5/3 CMJ (Ed)	EM480, 660, 670, 710, 740; RJ70-90, 110, HX Sections
OFREMP31	Employer Offers Health Ins RD 3/1 CMJ	EM690, 740; RJ100
OFREMP42	Employer Offers Health Ins RD 4/2 CMJ	EM690, 740; RJ100
OFREMP53	Employer Offers Health Ins RD 5/3 CMJ	EM690, 740; RJ100
EMPST31H	Employment Status RD 3/1 (Imp)	Full-Year Consolidated File: Missing Values Imputed
EMPST42H	Employment Status RD 4/2 (Imp)	Full-Year Consolidated File: Missing Values Imputed
EMPST53H	Employment Status RD 5/3 (Imp)	Full-Year Consolidated File: Missing Values Imputed
SLFCM31H	Self-employed at RD 3/1 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
SLFCM42H	Self-employed at RD 4/2 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
SLFCM53H	Self-employed at RD 5/3 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
NMEMP31H	Number of Employees at RD 3/1 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
NMEMP42H	Number of Employees at RD 4/2 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed

VARIABLE	DESCRIPTION	SOURCE
NMEMP53H	Number of Employees at RD 5/3 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
MORE31H	RD 3/1 CMJ Firm More Than 1 Locat (Imp)	Full-Year Consolidated File: Missing Values Imputed
MORE42H	RD 4/2 CMJ Firm More Than 1 Locat (Imp)	Full-Year Consolidated File: Missing Values Imputed
MORE53H	RD 5/3 CMJ Firm More Than 1 Locat (Imp)	Full-Year Consolidated File: Missing Values Imputed
INDCT31H	Industry Group RD 3/1 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
INDCT42H	Industry Group RD 4/2 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
INDCT53H	Industry Group RD 5/3 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
OCCCT31H	Occupation Group RD 3/1 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
OCCCT42H	Occupation Group RD 4/2 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
OCCCT53H	Occupation Group RD 5/3 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed

VARIABLE	DESCRIPTION	SOURCE
HOUR31H	Hours Per Week at RD 3/1 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
HOUR42H	Hours Per Week at RD 4/2 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
HOUR53H	Hours Per Week at RD 5/3 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
JBORG31H	Prv, St-Lc Gov, Fed Gov RD 3/1 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
JBORG42H	Prv, St-Lc Gov, Fed Gov RD 4/2 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
JBORG53H	Prv, St-Lc Gov, Fed Gov RD 5/3 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
UNION31H	Union Status at RD 3/1 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
UNION42H	Union Status at RD 4/2 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
UNION53H	Union Status at RD 5/3 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
BSNTY31H	Sol Prop, Prtnr, Corp, RD 3/1 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed

VARIABLE	DESCRIPTION	SOURCE
BSNTY42H	Sol Prop, Prtnr, Corp, RD 4/2 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
BSNTY53H	Sol Prop, Prtnr, Corp, RD 5/3 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
HRWG31H	Hourly Wage RD 3/1 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
HRWG42H	Hourly Wage RD 4/2 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
HRWG53H	Hourly Wage RD 5/3 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
CMJHLD31	Hlth Insur Held from RD 3/1 CMJ (PRPL)	PRPL PUF
CMJHLD42	Hlth Insur Held from RD 4/2 CMJ (PRPL)	PRPL PUF
CMJHLD53	Hlth Insur Held from RD 5/3 CMJ (PRPL)	PRPL PUF
OFFER31H	Health Insur Offered by RD 3/1 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
OFFER42H	Health Insur Offered by RD 4/2 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
OFFER53H	Health Insur Offered by RD 5/3 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
OFEMP31H	Emp Offers Health Ins RD 3/1 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed

VARIABLE	DESCRIPTION	SOURCE
OFEMP42H	Emp Offers Health Ins RD 4/2 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
OFEMP53H	Emp Offers Health Ins RD 5/3 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
PYVAC31H	Paid Vacation at RD 3/1 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
PYVAC42H	Paid Vacation at RD 4/2 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
PYVAC53H	Paid Vacation at RD 5/3 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
SCPAY31H	Paid Sick Leave at RD 3/1 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
SCPAY42H	Paid Sick Leave at RD 4/2 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
SCPAY53H	Paid Sick Leave at RD 5/3 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
PAYDR31H	Paid Leave to Visit Dr RD 3/1 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
PAYDR42H	Paid Leave to Visit Dr RD 4/2 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed

VARIABLE	DESCRIPTION	SOURCE
PAYDR53H	Paid Leave to Visit Dr RD 5/3 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
RTPLN31H	Pension Plan at RD 3/1 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
RTPLN42H	Pension Plan at RD 4/2 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed
RTPLN53H	Pension Plan at RD 5/3 CMJ (Imp)	Full-Year Consolidated File: Missing Values Imputed

HEALTH INSURANCE VARIABLES - PUBLIC USE

MONTHLY HEALTH INSURANCE COVERAGE INDICATORS

VARIABLE	DESCRIPTION	SOURCE
TRImm21X	Covered by TRICARE/CHAMPVA in mm 21 (Ed), where mm = JA-DE	HX120_01, 125_01, 230, PR260_01-280_01, HQ Section
MCRmm21	Covered by Medicare in mm 21, where mm = JA-DE	HX40_01, 50_01, 60, 230
MCRmm21X	Covered by Medicare in mm 21 (Ed), where mm = JA-DE	HX40_01, 50_01, 60, 230, see Section 2.5.11 for additional edit specifications
MCDmm21	Cov by Medicaid or SCHIP in mm 21, where mm = JA-DE	HX100, 230, PR150_01-160_01 and HQ Section
MCDmm21X	Cov by Medicaid or SCHIP in mm 21 (Ed), where mm = JA-DE	HX100_01, 230, PR150_01-160_01 and MCDmm20, Section 2.5.11 for additional edit specifications
GVAmm21	Cov by Other Public Coverage in mm 21, where mm = JA-DE	HQ Section, HX170_01, 230, PR310_01-320_01 and HQ Section
GVBmm21	Cov by Other Public HMO in mm 21, where mm = JA-DE	HX170_01, 230, 530-540, PR310_01-320_01, 340-350 and HQ Section

VARIABLE	DESCRIPTION	SOURCE
GVCmm21	Cov by Other Public Pays Prem in mm 21, where mm = JA-DE	HX170_01, 230, 560-570, PR310_01- 320_01, 370-380 and HQ Section
VAPmm21	Covered by VA in mm 21, where mm=JA-DE	HX120_01- 125_01, 230, PR260_01 - 280_01
IHSmm21	Cov by Indian Health Service in mm 21, Where mm=JA-DE	HX140, PR290_01, 300_01
PUBmm21X	Covr by Any Public Ins in mm 21 (Ed), where mm = JA-DE	TRImm21X, MCRmm21X, MCDmm21X, GVAm21, VAPmm21
PEGmm21	Covered by Empl Union Ins in mm 21, where mm = JA-DE	HX20, HX200, 300, 620; HP, OE, HQ, EM, RJ Sections
PDKmm21	Covr by Priv Ins (Source Unknwn) mm 21, where mm = JA-DE	HX200, 300, 620, HP, OE, and HQ Sections
PNGmm21	Covered by Nongroup Ins in mm 21, where mm = JA-DE	HX200, 300, 620, HP, OE, and HQ Sections
POGmm21	Covered by Other Group Ins in mm 21, where mm = JA-DE	HX200, 300, 620, HP, OE, and HQ Sections
POEmm21	Cov by ESI, Pholder Outside RU in mm 21, where mm = JA-DE	HX200, 300, 620, HP130, HQ, OE, RJ and EM sections
PNEmm21	Cov by Non-ESI, Phldr Outside RU in mm 21, where mm = JA-DE	HX200, 300, 620, HP130, OE, and HQ Sections
PRXmm21	Cov by Priv Ins through Exchng in mm 21, where mm = JA-DE	HX200, 300, 620,HP, OE, and HQ Sections

VARIABLE	DESCRIPTION	SOURCE
PRImm21	Covered by Private Ins in mm 21, where mm = JA-DE	POGmm21, PDKmm21, PEGmm21, PNEmm21, POEmm21, PNGmm21, PRXmm21
HPEmm21	Pholder of Empl Union Ins in mm 21, where mm = JA-DE	PEGmm21, HP70
HPDmm21	Pholder of Priv Ins (Source Unknwn) mm 21, where mm = JA-DE	PDKmm21; HP90
HPNmm21	Pholder of Nongroup Ins in mm 21, where mm = JA-DE	PNGmm21; HP90
HPOmm21	Pholder of Other Group Ins in mm 21, where mm = JA-DE	POGmm21; HP90
HPXmm21	Pholder of Priv Ins through Exch in mm 21, where mm = JA-DE	PRXmm21; HP90
HPRmm21	Pholder of Private Insurance in mm 21, where mm = JA-DE	HPEmm21, HPNmm21, HPDmm21, HPXmm21
INSmm21X	Covr by Hosp/Med Ins in mm 21 (Ed), where mm = JA-DE	PUBmm21X, PRImm21

SUMMARY HEALTH INSURANCE COVERAGE INDICATORS

VARIABLE	DESCRIPTION	SOURCE
PRVEV21	Ever Have Private Insurance during 2021	Constructed
TRIEV21	Ever Have TRICARE/CHAMPVA during 2021	Constructed
MCREV21	Ever Have Medicare during 2021 (ED)	Constructed
MCDEV21	Ever Have MCAID/SCHIP during 2021 (ED)	Constructed
VAEV21	Ever Have VA during 2021	Constructed
GVAEV21	Ever Have Other Public during 2021	Constructed
GVBEV21	Ever Have Other Public HMO during 2021	Constructed
GVCEV21	Ever Have Oth Pub Pays Prem during 2021	Constructed
UNINS21	Uninsured All of 2021	Constructed
INSCOV21	Health Insurance Coverage Indicator 2021	Constructed

VARIABLE	DESCRIPTION	SOURCE
INSURC21	Full Year Insurance Coverage Status 2021	Constructed

MANAGED CARE VARIABLES

VARIABLE	DESCRIPTION	SOURCE
TRI31X	Cov By TRICARE - Any Time in R3/1	HX125_01, 260, PR280_01
TRI42X	Cov By TRICARE - Any Time in R4/2	HX125_01, 260, PR280_01
TRI21X	Cov By TRICARE - R5/3 Til 12/31/21	HX125_01, 260, PR280_01
TRICH31X	Covered By CHAMPVA - Any Time in R3/1	HX230, 260, PR260_01, 270_01, 280_01, HQ Section
TRICH42X	Covered By CHAMPVA - Any Time in R4/2	HX230, 260, PR260_01, 270_01, 280_01, HQ Section
TRICH21X	Covered By CHAMPVA - R5/3 Until 12/31/21	HX230, 260, PR260_01, 270_01, 280_01, HQ Section
MCRPD31	Cov By Medicare Pmed Benefit - R3/1	HX40_01, 50_01, 60, 230, 410, 470, PR30, 90, HQ Section
MCRPD42	Cov By Medicare Pmed Benefit - R4/2	HX40_01, 50_01, 60, 230, 410, 470, PR30, 90, HQ Section
MCRPD21	Cov By Medicare Pmed Benefit - 12/31/21	HX40_01, 50_01, 60, 230, 410, 470, PR30, 90, HQ Section
MCRPD31X	Cov By Medicare Pmed Benefit - R3/1 (ED)	MCARE31X, MCAID31X, MCRPD31

VARIABLE	DESCRIPTION	SOURCE
MCRPD42X	Cov By Medicare Pmed Benefit - R4/2 (ED)	MCARE42X, MCAID42X, MCRPD42
MCRPD21X	Cov By Mcare Pmed Benefit-12/31/21 (ED)	MCARE21X, MCAID21X, MCRPD21
MCRPB31	Cov By Medicare Part B - R3/1	HX40_01, 50_01, 60, 230, 340, 350, and HQ section
MCRPB42	Cov By Medicare Part B - R4/2	HX40_01, 50_01, 60, 230, 340, 350, and HQ section
MCRPB21	Cov By Medicare Part B - 12/31/21	HX40_01, 50_01, 60, 230, 340, 350, and HQ section
MCRPHO31	Cov By Medicare Managed Care/Medicare Advantage - R3/1	HX40_01, 50_01, 60, 230, 390, PR10, HQ Section
MCRPHO42	Cov By Medicare Managed Care/Medicare Advantage - R4/2	HX40_01, 50_01, 60, 230, 390, PR10, HQ Section
MCRPHO21	Cov By Medicare Managed Care/Medicare Advantage - 12/31/21	HX40_01, 50_01, 60, 230, 390, PR10, HQ Section
MCDHMO31	Cov By Mcaid/SCHIP HMO -Any Time in R3/1	MCDAT31X, HX100, 230, 530 PR150_01- 160_01, 180, 340, and HQ Section
MCDHMO42	Cov By Mcaid/SCHIP HMO -Any Time in R4/2	MCDAT42X, HX100, 230, 530 PR150_01- 160_01, 180, 340, and HQ Section
MCDHMO21	Cov By Mcaid/SCHIP HMO-R5/3 Til 12/31/21	MCDAT21X, HX100, 230, 530 PR150_01- 160_01, 180, 340, and HQ Section

VARIABLE	DESCRIPTION	SOURCE
MCDMC31	Cv Mcaid/CHIP Gtkpr Pln-Any Time in R3/1	MCDHMO31, MCDAT31X, HX100, 230, 540 PR150_01- 160_01, 190, 350, and HQ Section
MCDMC42	Cv Mcaid/CHIP Gtkpr Pln-Any Time in R4/2	MCDHMO42, MCDAT42X, HX100, 230, 540, PR150_01- 160_01, 190, 350 and HQ Section
MCDMC21	Cv Mcd/CHIP Gtkpr Pln-R5/3 Til 12/31/21	MCDHMO21, MCDAT21X, HX100, 230, 540, PR150_01- 160_01, 190, 350 and HQ Section
PRVHMO31	Covered By Private HMO-Any Time in R3/1	MC10, HX200, 300, 620; HP, OE, HQ, EM, and RJ Sections
PRVHMO42	Covered By Private HMO-Any Time in R4/2	MC10, HX200, 300, 620; HP, OE, HQ, EM, and RJ Sections
PRVHMO21	Covered By Private HMO-R5/3 Til 12/31/21	MC10, HX200, 300, 620; HP, OE, HQ, EM, and RJ Sections

FLEXIBLE SPENDING ACCOUNT VARIABLES

VARIABLE	DESCRIPTION	SOURCE
FSAGT31	Anyone in RU Have FSA - R3/1	HX720
HASFSA31	Person is FSA Holder - R3/1	HX730
PFSAMT31	Person-level FSA Total Amount - R3/1	HX740_01

DURATION OF HEALTH INSURANCE VARIABLES

VARIABLE	DESCRIPTION	SOURCE
PREVCOVR	Per Cov by Ins in Prev 2 Yrs-Pnl 25 Only	HX750
MORECOVR	Cov by Mor Compr Pl Prev 2 Yr-Pn 25 Only	HX760

OTHER HEALTH INSURANCE COVERAGE VARIABLES

VARIABLE	DESCRIPTION	SOURCE
TRICR31X	Cov by TRICR/CHAMV - R3/1 Int Dt (Ed)	Constructed
TRICR42X	Cov by TRICR/CHAMV - R4/2 Int Dt (Ed)	Constructed
TRICR53X	Cov by TRICR/CHAMV 12-31/R3 Int Dt (Ed)	Constructed
TRICR21X	Cov by TRICR/CHAMV - 12/31/21 (Ed)	Constructed
TRIAT31X	Cov TRICARE/CHAMPVA - Any Time in R3/1	Constructed
TRIAT42X	Cov TRICARE/CHAMPVA - Any Time in R4/2	Constructed
TRIAT53X	Cov TRICARE/CHAMPVA - Any Time in R5/3	Constructed
TRIAT21X	Cov TRICARE/CHAMPVA-R5/3 Until 12/31/21	Constructed
MCAID31	Cov by Medicaid or SCHIP - R3/1 Int Dt	Constructed
MCAID42	Cov by Medicaid or SCHIP - R4/2 Int Dt	Constructed
MCAID53	Cov by Medicaid or SCHIP 12-31/R3 Int Dt	Constructed
MCAID21	Cov by Medicaid or SCHIP - 12/31/21	Constructed
MCAID31X	Cov by Medicaid/SCHIP - R3/1 Int Dt (Ed)	Constructed
MCAID42X	Cov by Medicaid/SCHIP - R4/2 Int Dt (Ed)	Constructed
MCAID53X	Cov Medicaid/SCHIP 12-31/R3 Int Dt (Ed)	Constructed
MCAID21X	Cov by Medicaid or SCHIP - 12/31/21 (Ed)	Constructed
MCARE31	Cov by Medicare - R3/1 Int Dt	Constructed
MCARE42	Cov by Medicare - R4/2 Int Dt	Constructed
MCARE53	Cov by Medicare 12-31/R3 Int Dt	Constructed
MCARE21	Cov by Medicare - 12/31/21	Constructed
MCARE31X	Cov by Medicare - R3/1 Int Dt (Ed)	Constructed
MCARE42X	Cov by Medicare - R4/2 Int Dt (Ed)	Constructed
MCARE53X	Cov by Medicare 12-31/R3 Int Dt (Ed)	Constructed
MCARE21X	Cov by Medicare - 12/31/21 (Ed)	Constructed
MCDAT31X	Cov Medicaid Or SCHIP - Any Time in R3/1	Constructed

VARIABLE	DESCRIPTION	SOURCE
MCDAT42X	Cov Medicaid Or SCHIP - Any Time in R4/2	Constructed
MCDAT53X	Cov Medicaid Or SCHIP - Any Time in R5/3	Constructed
MCDAT21X	Cv Mcd/CHIP-Any Tme In R5/3 Til 12/31/21	Constructed
GOVTA31	Cov by Other Public Cov - R3/1 Int Dt	Constructed
GOVTA42	Cov by Other Public Cov - R4/2 Int Dt	Constructed
GOVTA53	Cov by Other Public Cov 12-31/R3 Int Dt	Constructed
GOVTA21	Cov by Other Public Cov - 12/31/21	Constructed
GOVAAT31	Cov by Other Public -Any Time in R3/1	Constructed
GOVAAT42	Cov by Other Public -Any Time in R4/2	Constructed
GOVAAT53	Cov by Other Public -Any Time in R5/3	Constructed
GOVAAT21	Cv Oth Pub- Any Time in R5/3 Til 12/31/21	Constructed
GOVTB31	Cov By Oth Pub HMO - R3/1 Int Dt	Constructed
GOVTB42	Cov By Oth Pub HMO - R4/2 Int Dt	Constructed
GOVTB53	Cov By Oth Pub HMO 12-31/R3 Int Dt	Constructed
GOVTB21	Cov By Oth Pub HMO - 12/31/21	Constructed
GOVBAT31	Cov By Oth Pub HMO-Any Time in R3/1	Constructed
GOVBAT42	Cov By Oth Pub HMO-Any Time in R4/2	Constructed
GOVBAT53	Cov By Oth Pub HMO-Any Time in R5/3	Constructed
GOVBAT21	Cov Oth Pb HMO-Any Tm R5/3 Til 12/31/21	Constructed
GOVTC31	Cov By Oth Pub Pays Prem - R3/1 Int Dt	Constructed
GOVTC42	Cov By Oth Pub Pays Prem - R4/2 Int Dt	Constructed
GOVTC53	Cov By Oth Pub Pays Prem 12-31/R3 Int Dt	Constructed
GOVTC21	Cov By Oth Pub Pays Prem - 12/31/21	Constructed
GOVCAT31	Cv By Oth Pub Pays Prem-Any Time in R3/1	Constructed
GOVCAT42	Cv By Oth Pub Pays Prem-Any Time in R4/2	Constructed
GOVCAT53	Cv By Oth Pub Pays Prem-Any Time in R5/3	Constructed
GOVCAT21	Cv Oth Pub Pays Prem- R5/3 Til 12/31/21	Constructed
VAPROG31	Cov By VA - R3/1 Int Dt	Constructed
VAPROG42	Cov By VA - R4/2 Int Dt	Constructed
VAPROG53	Cov By VA 12-31/R3 Int Dt	Constructed
VAPROG21	Cov By VA- 12/31/21	Constructed
VAPRAT31	Cov By VA - Any Time in R3/1	Constructed
VAPRAT42	Cov By VA - Any Time in R4/2	Constructed

VARIABLE	DESCRIPTION	SOURCE
VAPRAT53	Cov By VA - Any Time in R5/3	Constructed
VAPRAT21	Cov By VA-Any Tme In R5/3 Til 12/31/21	Constructed
IHS31	Cv By Indian Health Service- R3/1 Int Dt	Constructed
IHS42	Cv By Indian Health Service- R4/2 Int Dt	Constructed
IHS53	Cv By IHS 12-31/R3 Int Dt	Constructed
IHS21	Cov By IHS - 12/31/21	Constructed
IHSAT31	Cov By IHS - Any Time in R3/1	Constructed
IHSAT42	Cov By IHS - Any Time in R4/2	Constructed
IHSAT53	Cov By IHS - Any Time in R5/3	Constructed
IHSAT21	Cov By IHS-Any Tme In R5/3 Til 12/31/21	Constructed
PRIDK31	Cov by Priv Ins (Dk Plan) - R3/1 Int	Constructed
PRIDK42	Cov by Priv Ins (Dk Plan) - R4/2 Int	Constructed
PRIDK53	Cov by Priv Ins (Dk Plan) 12-31/R3 Int	Constructed
PRIDK21	Cov by Priv Ins (Dk Plan) - 12/31/21	Constructed
PRIEU31	Cov by Empl/Union Grp Ins - R3/1 Int Dt	Constructed
PRIEU42	Cov by Empl/Union Grp Ins - R4/2 Int Dt	Constructed
PRIEU53	Cov by Empl/Union Grp Ins 12-31/R3 Int	Constructed
PRIEU21	Cov by Empl/Union Grp Ins - 12/31/21	Constructed
PRING31	Cov by Non-Group Ins - R3/1 Int Dt	Constructed
PRING42	Cov by Non-Group Ins - R4/2 Int Dt	Constructed
PRING53	Cov by Non-Group Ins 12-31/R3 Int Dt	Constructed
PRING21	Cov by Non-Group Ins - 12/31/21	Constructed
PRIOG31	Cov by Other Group Ins - R3/1 Int Dt	Constructed
PRIOG42	Cov by Other Group Ins - R4/2 Int Dt	Constructed
PRIOG53	Cov by Other Group Ins 12-31/R3 Int Dt	Constructed
PRIOG21	Cov by Other Group Ins - 12/31/21	Constructed
PRINEO31	Cov By Non-ESI, PHoldr Out RU - R3/1 Int Dt	Constructed
PRINEO42	Cov By Non-ESI, PHoldr Out RU - R4/2 Int Dt	Constructed
PRINEO53	Cov By Non-ESI, PHoldr Out RU 12-31/R3 Int Dt	Constructed
PRINEO21	Cov By Non-ESI, PHoldr Out RU- 12/31/21	Constructed
PRIEUO31	Cov By ESI, Pholdr Outside RU - R3/1 Int	Constructed
PRIEUO42	Cov By ESI, Pholdr Outside RU - R4/2 Int	Constructed
PRIEUO53	Cv By ESI, Phldr Outside RU 12-31/R3 Int	Constructed

VARIABLE	DESCRIPTION	SOURCE
PRIEUO21	Cov By ESI, PHoldr Outside RU - 12/31/21	Constructed
PRSTX31	Cov by Priv Exchange Ins - R3/1 Int Dt	Constructed
PRSTX42	Cov by Priv Exchange Ins - R4/2 Int Dt	Constructed
PRSTX53	Priv Exchange Ins on 12-31 R5/R3 Int Dt	Constructed
PRSTX21	Priv Exchange Insurance on 12/31/21	Constructed
PRIV31	Cov by Priv Hlth Ins - R3/1 Int Date	Constructed
PRIV42	Cov by Priv Hlth Ins - R4/2 Int Date	Constructed
PRIV53	Cov by Priv Hlth Ins 12-31/R3 Int Date	Constructed
PRIV21	Cov by Priv Hlth Ins - 12/31/21	Constructed
PRIVAT31	Cov by Private Ins - Any Time in R3/1	Constructed
PRIVAT42	Cov by Private Ins - Any Time in R4/2	Constructed
PRIVAT53	Cov by Private Ins - Any Time in R5/3	Constructed
PRIVAT21	Cov by Private Ins - R5/3 Until 12/31/21	Constructed
PUB31X	Cov by Public Ins - R3/1 Int Dt (Ed)	Constructed
PUB42X	Cov by Public Ins - R4/2 Int Dt (Ed)	Constructed
PUB53X	Cov by Public Ins 12-31/R3 Int Dt (Ed)	Constructed
PUB21X	Cov by Public Ins - 12/31/21 (Ed)	Constructed
PUBAT31X	Cov by Public - Any Time in R3/1	Constructed
PUBAT42X	Cov by Public - Any Time in R4/2	Constructed
PUBAT53X	Cov by Public - Any Time in R5/3	Constructed
PUBAT21X	Cov by Public - R5/3 Until 12/31/21	Constructed
VERFLG31	PID Gained Covg Thru Verification R3/1	Constructed
VERFLG42	PID Gained Covg Thru Verification R4/2	Constructed
VERFLG21	PID Gained Covg Thru Verification in R5/R3 Until 12/31/21	Constructed
INS31X	Insured - R3/1 Int Date (Ed)	Constructed
INS42X	Insured - R4/2 Int Date (Ed)	Constructed
INS53X	Insured 12-31/R3 Int Date (Ed)	Constructed
INS21X	Insured - 12/31/21 (Ed)	Constructed
INSAT31X	Insured Any Time in R3/1	Constructed
INSAT42X	Insured Any Time in R4/2	Constructed
INSAT53X	Insured Any Time in R5/3	Constructed
INSAT21X	Insured Any Time in R5/R3 until 12/31/21	Constructed

DENTAL AND PRESCRIPTION DRUG PRIVATE INSURANCE VARIABLES

VARIABLE	DESCRIPTION	SOURCE
DENTIN31	Dental Insurance- Any Time in RD 3/1	Constructed
DENTIN42	Dental Insurance- Any Time in RD 4/2	Constructed
DENTIN53	Dental Insurance- Any Time in RD 5/3	Constructed
DNTINS31	Dental Ins - Any Time In Rd 3/1 in 2021	Constructed
DNTINS21	Dentl Ins-Any Time in R5/R3 Til 12/31/21	Constructed
PMEDIN31	PMED Ins - Any Time in Rd 3/1	Constructed
PMEDIN42	PMEDINS - Any Time in Rd 4/2	Constructed
PMEDIN53	PMEDINS - Any Time in Rd 5/3	Constructed
PMDINS31	Pmed Ins - RD 3/1 in 2021	Constructed
PMDINS21	Pmed Ins - R5/R3 until 12/31/21	Constructed

MEDICAL DEBT VARIABLES

VARIABLE	DESCRIPTION	SOURCE
PROBPY42	Family Having Prob Paying Medical Bills	HX770
CRFMPY42	Family Med Bills Being Paid Over Time	HX780
PYUNBL42	Unable to Pay Family Medical Bills	HX790

THIRD PARTY PAYER VARIABLES - PUBLIC USE

VARIABLE	DESCRIPTION	SOURCE
PMEDUP31	Has Usual 3rd Party Payer for Pmeds - R3/1	Constructed
PMEDUP42	Has Usual 3rd Party Payer for Pmeds - R4/2	Constructed
PMEDUP53	Has Usual 3rd Party Payer for Pmeds - R5/3	Constructed
PMEDPY31	Usual 3rd Party Payer for Pmeds - R3/1	Constructed
PMEDPY42	Usual 3rd Party Payer for Pmeds - R4/2	Constructed
PMEDPY53	Usual 3rd Party Payer for Pmeds - R5/3	Constructed

PERSON-LEVEL UTILIZATION VARIABLES - PUBLIC USE

VARIABLE	DESCRIPTION	SOURCE
OBTOTV21	# Office-Based Provider Visits 2021	Constructed
OBDRV21	# Office-Based Physician Visits 2021	Constructed
OPTOTV21	# Outpatient Dept Provider Visits 2021	Constructed
OPDRV21	# Outpatient Dept Physician Visits 2021	Constructed
ERTOT21	# Emergency Room Visits 2021	Constructed
IPDIS21	# Hospital Discharges 2021	Constructed
IPNGTD21	# Nights in Hosp for Discharges 2021	Constructed
DVTOT21	# Dental Care Visits 2021	Constructed
HHTOTD21	# Home Health Provider Days 2021	Constructed
HHAGD21	# Agency Home Health Provider Days 2021	Constructed
HHINDD21	# Non-Agency Home Hlth Providr Days 2021	Constructed
HHINFD21	# Informal Home Hlth Provider Days 2021	Constructed
RXTOT21	# Prescribed Medicines including Refills 2021	Constructed

WEIGHTS VARIABLES - PUBLIC USE

VARIABLE	DESCRIPTION	SOURCE
PERWT21F	Final Person Weight, 2021	Constructed
FAMWT21F	Final Family Weight, 2021	Constructed
FAMWT21C	Pov Adj Family Weight-CPS Fam on 12/31/21	Constructed
SAQWT21F	Final SAQ Person Weight, 2021	Constructed
DIABW21F	Final Diabetes Care Supplement Weight	Constructed
SDOHWT21F	Final SDOH Weight, 2021	Constructed
VARSTR	Variance Estimation Stratum - 2021	Constructed
VARPSU	Variance Estimation PSU - 2021	Constructed

Appendix 1

MEPS Industry Codes Condensing Rules

MEPS Industry Codes Condensing Rules FY2010 and Subsequent Files

Condensed Industry Code	Census Industry Code Range	Description
1	0170 - 0290	Natural Resources
2	0370 - 0490	Mining
3	0770	Construction
4	1070 - 3990	Manufacturing
5	4070 - 5790	Wholesale and Retail Trade
6	0570 - 0690, 6070 - 6390	Transportation and Utilities
7	6470 - 6780	Information
8	6870 - 7190	Financial Activities
9	7270 - 7790	Professional and Business Services
10	7860 - 8470	Education, Health, and Social Services
11	8560 - 8690	Leisure and Hospitality
12	8770 - 9290	Other Services
13	9370 - 9590	Public Administration
14	9890	Military
15	9990	Unclassifiable Industry

MEPS uses the 4-digit Census occupation and industry coding systems developed for the Current Population Survey and the American Community Survey.

For industry coding, MEPS uses the 2007 4-digit Census industry codes. Descriptions of the 4-digit Census industry codes can be found at the [U.S. Bureau of Labor Statistics website](http://www.bls.gov).

See [Census IO Index](#) for more information on the Census coding systems used by MEPS.

Appendix 2

MEPS Occupation Codes Condensing Rules

MEPS Occupation Codes Condensing Rules FY2010 and Subsequent Files

Condensed Occupation Code	Census Occupation Code Range	Description
1	0010 - 0950	Management, Business, and Financial Operations Occupations
2	1005 - 3540	Professional and Related Occupations
3	3600 - 4650	Service Occupations
4	4700 - 4965	Sales and Related Occupations
5	5000 - 5940	Office and Administrative Support Occupations
6	6005 - 6130	Farming, Fishing, and Forestry Occupations
7	6200 - 7630	Construction, Extraction, and Maintenance Occupations
8	7700 - 9750	Production, Transportation, and Material Moving Occupations
9	9840	Military Specific Occupations
10	9920	Not in Labor Force
11	9990	Unclassifiable Occupation

MEPS uses the 4-digit Census occupation and industry coding systems developed for the Current Population Survey and the American Community Survey.

For occupation coding, MEPS uses the 2010 4-digit Census occupation codes. Descriptions of the 4-digit Census occupation codes can be found at the [U.S. Bureau of Labor Statistics website](#).

See the [Census IO Index](#) for more information on the Census coding systems used by MEPS.

Appendix 3

Summary of Utilization and Expenditure Variables by Health Service Category

HEALTH SERVICE CATEGORY	UTILIZATION VARIABLE(S)	EXPENDITURE VARIABLE(S) ¹
<i>All Health Services</i>	--	TOT***21
Total Office Based Visits (Physician + Non-physician + Unknown)	OBTOTV21	OBV***21
Office Based Visits to Physicians	OBDRV21	OBD***21
Total Outpatient Visits (Physician + Non-physician + Unknown)	OPTOTV21	--
Sum of Facility and SBD Expenses	--	OPT***21
Facility Expense	--	OPF***21
SBD Expense	--	OPD***21
Outpatient Visits to Physicians	OPDRV21	--
Facility Expense	--	OPV***21
SBD Expense	--	OPS***21
Total Emergency Room Visits	ERTOT21	--
Sum of Facility and SBD Expenses	--	ERT***21
Facility Expense	--	ERF***21
SBD Expense	--	ERD***21
Total Inpatient Stays	IPDIS21, IPNGTD21	--
Sum of Facility and SBD Expenses	--	IPT***21
Facility Expense	--	IPF***21
SBD Expense	--	IPD***21
Total Prescription Medicines	RXTOT21	RX***21
Total Dental Visits	DVTOT21	DVT***21
Total Home Health Care	HHTOTD21	--
Agency Sponsored	HHAGD21	HHA***21
Paid Independent Providers	HHINDD21	HHN***21

¹ See key at end of table for specific categories for ***.

HEALTH SERVICE CATEGORY	UTILIZATION VARIABLE(S)	EXPENDITURE VARIABLE(S)¹
Informal	HHINFD21	--
Vision Aids	--	VIS***21
Other Medical Supplies and Equipment	--	OTH***21

KEY: To complete variable name, replace *** with a particular source of payment category as identified in the following tables:

Source of Payment Category	***
Total payments (sum of all sources)	EXP
Out of Pocket	SLF
Medicare	MCR
Medicaid	MCD
Private Insurance	PRV
Veteran's Administration/CHAMPVA	VA
TRICARE	TRI
Other Federal Sources	OFD
Other State and Local Sources	STL
Workers' Compensation	WCP
Other Unclassified Sources	OSR

Collapsed Source of Payment Category	***
Private and TRICARE	PTR
Other Federal, Other State and Local, and Other Unclassified Sources	OTH
Total charges²	TCH

² No charge variables on file for prescription medicines.