MEPS HC-245
Panel 24, 4-Year Longitudinal Public Use File

November 2024

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

Table of Contents

Section			Page
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-1
		2.1 Variables	C-2
		2.1.1 Variables from Annual Full-Year Consolidated Files	C-2
		2.1.2 Constructed Variables for Selection of Group	C-8
		2.1.3 Estimation Variables	C-9

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:

> No one is to use the data in this data set in any way except for statistical reporting and analysis; and
>
> 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
>
> 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 non-institutionalized 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 healthcare. Estimates can be produced for individuals, families, and selected population subgroups. The panel design of the survey includes 5 rounds of interviews covering 2 full calendar years. Additional rounds were added to Panel 24 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. Information about each household member is collected through computer-assisted personal interviewing (CAPI) technology, 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. Historically, each annual MEPS-HC sample size consists of approximately up to 15,000 households. Data can be analyzed at the person, the family, 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 (NCHS). The NHIS sampling frame provides a nationally representative sample of the U.S. civilian noninstitutionalized population. In 2006, the NCHS implemented a new sample design for the NHIS, to include households with Asian persons in addition to households with Black and Hispanic persons in the oversampling of minority populations. In 2016, NCHS introduced another sample design that discontinued the oversampling of these minority groups.

2.0 Medical Provider Component

When the household CAPI interview is completed, and permission is obtained from the household survey respondents to contact their medical provider(s), a sample of these providers is contacted by telephone to obtain information that household sample members cannot accurately provide. This part of the MEPS is called the Medical Provider Component (MPC) and it collects information on dates of visit, diagnosis and procedure codes, and charges and payments. The Pharmacy Component (PC), a subcomponent of the MPC, does not collect data on charges or on diagnosis and procedure codes, but it does collect detailed information on drugs, including the 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. The MEPS HC data are collected under contract with Westat, Inc. and the MEPS MPC data are collected under contract with Research Triangle Institute. 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 NCHS provides consultation and technical assistance.

As soon as the MEPS data are collected and edited, they are released to the public in stages 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 Panel 24, 4-year Longitudinal Public Use File (PUF) from the MEPS-HC. It was released as an ASCII file (with related SAS, Stata, SPSS, and R programming statements and data user information) and as a SAS data set, a SAS transport dataset, a Stata dataset, and an Excel file. The Panel 24, 4-year Longitudinal PUF provides information collected from a nationally representative sample of the U.S. civilian noninstitutionalized population for the four-year period 2019-2022. The file contains 5,321 variables and has a logical record length of 15,038 with an additional 2-byte carriage return/line feed at the end of each record.

This file consists of MEPS survey data obtained in Rounds 1-9 of MEPS Panel 24 and can be used to analyze changes over a four-year period. 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 2019, 2020, 2021, and 2022 Full-Year Consolidated Files (HC-216, HC-224, HC-233, and HC-243, respectively).

The following documentation offers a brief overview of the contents and structure of the files and programming information. A codebook of all the variables included in the Panel 24, 4-year Longitudinal PUF is provided in a separate file (H245CB.PDF). A database of all MEPS products released to date and a variable locator indicating the major MEPS data items on public use files that have been released to date can be found on the MEPS website.

2.0 Data File Information

This Panel 24, 4-year Longitudinal PUF contains records for 5,565 persons in Panel 24 who were respondents for the period they were in-scope for the survey (i.e., a member of the civilian non-institutionalized population) during the four-year period. Only persons with positive person-level weights (PERWT19F, PERWT20F, PERWT21F, or PERWT22F) are included in the longitudinal PUF data. Data are available for all nine rounds for 87.74% of the cases (4,883 persons). The remaining 12.26% (682 persons) do not have data for one or more rounds but were in-scope for all rounds they participated in the survey. These persons are those who were born, died, were in the military or an institution, or left the country during the four-year period. In contrast, persons in the panel who participated in the survey for only part of the period they were in-scope are not included in this file. To compensate for this attrition, adjustments were made in the construction of the panel weight variable included in this file (LONGWT). The codebook provides both weighted and unweighted frequencies for each variable on the data file. The LONGWT variable should be used to produce national estimates for the four-year period.

2.1 Variables

2.1.1 Variables from Annual Full-Year Consolidated Files

Most variables on this file were obtained from the MEPS 2019, 2020, 2021, and 2022 Full-Year Consolidated Files (HC-216, HC-224, HC-233, and HC-243, respectively). A variable named YEARIND was added to identify the years of data in which each sample person is included. YEARIND may take the following values for Panel 24:

1 = All four years
2 = 2019 only
3 = 2020 only
4 = 2021 only
5 = 2022 only
6 = 2019 and 2020 only
7 = 2019 and 2021 only
8 = 2019 and 2022 only
9 = 2020 and 2021 only
10 = 2020 and 2022 only
11= 2021 and 2022 only
12 = 2019, 2020, and 2021 only
13 = 2019, 2020, and 2022 only
14 = 2019, 2021, and 2022 only
15 = 2020, 2021, and 2022 only

Names for time dependent variables from the full year files were modified in order to: 1) eliminate duplicate variable names for data reflecting different time periods during the panel, and 2) standardize variable names to facilitate pooling of multiple MEPS panels for analysis.[1] Generally, annual variables with a suffix of "19","20", "21", and "22" are renamed with a suffix of "Y1", "Y2", "Y3", and "Y4", respectively. Variables with a suffix of "31", "42", and "53" are renamed with a suffix denoting the round the data was collected (i.e., "1" , "2" or "3" for variables originating from Rounds 1-3 on the 2019 full-year file, "3", "4", or "5" for variables originating from Rounds 3-5 on the 2020 full-year file, "5", "6", or "7" for variables originating from Rounds 5-7 on the 2021 full-year file, and "7","8", or "9" for variables originating from Rounds 7-9 on the 2022 full-year file).[2] It is necessary to use this crosswalk in conjunction with documentation for the 2019, 2020, 2021, and 2022 full-year consolidated files to obtain a full description of variables on this file. Table 1 below provides the crosswalk summarizing the scheme used for renaming variables from the annual files.

Table 1. Crosswalk of Variable Names between the Full-Year Consolidated Files and the Panel 24, 4 Year Longitudinal File

[1] A variable named PANEL is also included to facilitate pooling across panels. This variable is simply the panel number and is therefore constant across all records within a longitudinal file. The ten-character variable DUPERSID uniquely identifies each person represented on the file and is the combination of the variables DUID (PANEL + Dwelling Unit ID) and PID (Person Number).

[2] While Round 3 values were obtained for most observations from the 2020 Full Year Consolidated File, they were obtained from the 2019 Full Year Consolidated File for sample persons not in the 2020 data (YEARIND=2, 7, 8, or 14). Similarly, values for Round 5 variables were obtained for most observations from the 2021 Full Year Consolidated File, but were obtained from the 2020 Full Year Consolidated File for sample persons not in the 2021 data (YEARIND=3, 6, 10, or 13). Round 7 values were obtained for most observations from the 2022 Full Year Consolidated file but were obtained from the 2021 Full Year Consolidated file for sample persons not in the 2022 data (YEARIND=4, 7, 9, or 12).

Type of Variable	Full-Year Consolidated PUF Variable Name Suffix	Longitudinal PUF Variable Name Suffix	Specific cases or examples
Constant (i.e., not round or year specific)	No suffixes	No suffixes	All variables: BORNUSA=BORNUSA DOBMM=DOBMM DOBYY=DOBYY DUID=DUID PID=PID DUPERSID=DUPERSID EDUCYR=EDUCYR HIDEG=HIDEG HISPANX=HISPANX HISPNCAT=HISPNCAT HWELLSPK=HWELLSPK INTVLANG=INTVLANG OTHLGSPK=OTHLGSPK PANEL=PANEL PID=PID RACEAX=RACEAX RACEBX=RACEBX RACEWX=RACEWX RACEV1X=RACEV1X RACEV2X=RACEV2X RACETHX=RACETHX SEX=SEX VARPSU=VARPSU VARSTR=VARSTR WHTLGSPK=WHTLGSPK YRSINUS=YRSINUS
Annual, family related variables	YR	Y1 or YR1	All variables: FAMIDYR=FAMIDYR1 (2019 file) FAMRFPYR=FAMRFPY1 (2019 file) FAMSZEYR=FAMSZYR1 (2019 file)
		Y2 or YR2	FAMIDYR=FAMIDYR2 (2020 file) FAMRFPYR=FAMRFPY2 (2020 file) FAMSZEYR=FAMSZYR2 (2020 file)
		Y3 or YR3	FAMIDYR=FAMIDYR3 (2021 file) FAMRFPYR=FAMRFPY3 (2021 file) FAMSZEYR=FAMSZYR3 (2021 file)
		Y4 or YR4	FAMIDYR=FAMIDYR4 (2022 file) FAMRFPYR=FAMRFPY4 (2022 file) FAMSZEYR=FAMSZYR4 (2022 file)

Type of Variable	Full-Year Consolidated PUF Variable Name Suffix	Longitudinal PUF Variable Name Suffix	Specific cases or examples
Annual, CPS family identifiers	No suffix	Y1	All variables: CPSFAMID= CPSFAMY1 (2019 file)
		Y2	CPSFAMID= CPSFAMY2 (2020 file)
		Y3	CPSFAMID= CPSFAMY3 (2021 file)
		Y4	CPSFAMID= CPSFAMY4 (2022 file)
Annual, health insurance eligibility units	No suffix	Y1	All variables: HIEUIDX=HIEUIDY1 (2019 file)
		Y2	HIEUIDX=HIEUIDY2 (2020 file)
		Y3	HIEUIDX=HIEUIDY3 (2021 file)
		Y4	HIEUIDX=HIEUIDY4 (2022 file)
Annual, in-scope variables	No suffixes	YR1	All variables: INSCOPE=INSCPYR1 (2019 file)
		YR2	INSCOPE=INSCPYR2 (2020 file)
		YR3	INSCOPE=INSCPYR3 (2021 file)
		YR4	INSCOPE=INSCPYR4 (2022 file)
12/31 status variables	1231 in 2019 file	Y1	All variables: FAMS1231=FAMSY1 (2019 file) FCRP1231=FCRPY1 (2019 file) FCSZ1231= FCSZY1 (2019 file) FMRS1231= FMRSY1 (2019 file) INSC1231=INSCY1 (2019 file)
	1231 in 2020 file	Y2	FAMS1231=FAMSY2 (2020 file) FCRP1231=FCRPY2 (2020 file) FCSZ1231= FCSZY2 (2020 file) FMRS1231= FMRSY2 (2020 file) INSC1231=INSCY2 (2020 file)
	1231 in 2021 file	Y3	FAMS1231=FAMSY3 (2021 file) FCRP1231=FCRPY3 (2021 file) FCSZ1231= FCSZY3 (2021 file) FMRS1231= FMRSY3 (2021 file) INSC1231=INSCY3 (2021 file)
	1231 in 2022 file	Y4	FAMS1231=FAMSY4 (2022 file) FCRP1231=FCRPY4 (2022 file) FCSZ1231= FCSZY4 (2022 file) FMRS1231= FMRSY4 (2022 file) INSC1231=INSCY4 (2022 file)

Type of Variable	Full-Year Consolidated PUF Variable Name Suffix	Longitudinal PUF Variable Name Suffix	Specific cases or examples
Annual	19, 19X, 19F, or 19C	Y1, Y1X, Y1F, or Y1C	Examples: TOTEXP19=TOTEXPY1 AGE19X=AGEY1X
	20, 20X, 20F, or 20C	Y2, Y2X, Y2F, or Y2C	TOTEXP20=TOTEXPY2 AGE20X=AGEY2X
	21, 21X, 21F, or 21C	Y3, Y3X, Y3F, or Y3C	TOTEXP21=TOTEXPY3 AGE21X=AGEY3X
	22, 22X, 22F, or 22C	Y4, Y4X, Y4F, or Y4C	TOTEXP22=TOTEXPY4 AGE22X=AGEY4X
Variables for health insurance prior to January 1, 2019 (data collected in Round 1 only)	No suffixes	No suffixes	All variables: PREVCOVR=PREVCOVR MORECOVR=MORECOVR

Type of Variable	Full-Year Consolidated PUF Variable Name Suffix	Longitudinal PUF Variable Name Suffix	Specific cases or examples
Annual	No suffixes[3]	Y1	Examples: KEYNESS=KEYNESY1 (2019 file) SAQELIG=SAQELIY1 (2019 file) EVRWRK=EVRWRKY1 (2019 file) EVRETIRE=EVRETIY1 (2019 file) AGELAST=AGELSTY1 (2019 file) DIABDX_M18=DIABDXY1_M18 (2019 file)
		Y2	KEYNESS=KEYNESY2 (2020 file) SAQELIG=SAQELIY2 (2020 file) EVRWRK=EVRWRKY2 (2020 file) EVRETIRE=EVRETIY2 (2020 file) AGELAST=AGELSTY2 (2020 file) DIABDX_M18=DIABDXY2_M18 (2020 file)
		Y3	KEYNESS=KEYNESY3 (2021 file) SAQELIG=SAQELIY3 (2021 file) EVRWRK=EVRWRKY3 (2021 file) EVRETIRE=EVRETIY3 (2021 file) AGELAST=AGELSTY3 (2021 file) DIABDX_M18=DIABDXY3_M18 (2021 file)
		Y4	KEYNESS=KEYNESY4 (2022 file) SAQELIG=SAQELIY4 (2022 file) EVRWRK=EVRWRKY4 (2022 file) EVRETIRE=EVRETIY4 (2022 file) AGELAST=AGELSTY4 (2022 file) DIABDX_M18=DIABDXY4_M18 (2022 file)
Monthly	2-character month + 19 2-character month + 20 2-character month + 21 2-character month + 22	2-character month + Y1 2-character month + Y2 2-character month + Y3 2-character month + Y4	Example: PRIJA18=PRIJAY1 (2019 file) PRIJA19=PRIJAY2 (2020 file) PRIJA20=PRIJAY3 (2021 file) PRIJA21=PRIJAY4 (2022 file)

[3] To maintain a previously-implemented 8-character naming convention, some variable names had the last character or two dropped in the renaming process. A few variables have names longer than 8 characters because they were modified and tagged with an '*_Myy'* suffix, where *yy* indicates the year of modification. These variables were altered in the same fashion they would have been without the *_Myy* suffix, and the *_Myy* suffix was retained.

Type of Variable	Full-Year Consolidated PUF Variable Name Suffix	Longitudinal PUF Variable Name Suffix	Specific cases or examples
			Examples:
Round Specific	31, 31X, or 31H in 2019 file	1, 1X, or 1H for 2019	RTHLTH31=RTHLTH1 (2019 file)
	42, 42X, or 42H in 2019 file	2, 2X, or 2H for 2019	RTHLTH42=RTHLTH2 (2019 file)
	53, 53X, or 53H in 2019 file	3, 3X, or 3H for 2019	RTHLTH53=RTHLTH3 (2019 file sample person is not in 2020 data)
	31_M18 in 2019 file	1_M18 for 2019	JTPAIN31_M18=JTPAIN1_M18
	42_M18 in 2019 file	2_M18 for 2019	PROVTY42_M18=PROVTY2_M18
	31, 31X, or 31H in 2020 file	3, 3X, 3H for 2020	RTHLTH31= RTHLTH3 (2020 file if sample person is in 2020 data)
	42, 42X, or 42H in 2020 file	4, 4X, 4H for 2020	RTHLTH42=RTHLTH4 (2020 file)
	53, 53X, or 53H in 2020 file	5, 5X, 5H for 2020	RTHLTH53=RTHLTH5 (2020 file if sample person is not in 2021 data)
	31_M18 in 2020 file	3_M18 for 2020	JTPAIN31_M18=JTPAIN3_M18
	42_M18 in 2020 file	4_M18 for 2020	PROVTY42_M18=PROVTY4_M18
	31, 31X, or 31H in 2021 file	5, 5X, 5H for 2021	RTHLTH31= RTHLTH5 (2021 file if sample person is in 2021 data)
	42, 42X, or 42H in 2021 file	6, 6X, 6H for 2021	RTHLTH42=RTHLTH6 (2021 file)
	53, 53X, or 53H in 2021 file	7, 7X, 7H for 2021	RTHLTH53=RTHLTH7 (2021 file if sample person is not in 2022 data)
	31_M18 in 2021 file	5_M18 for 2021	JTPAIN31_M18=JTPAIN5_M18
	42_M18 in 2021 file	6_M18 for 2021	PROVTY42_M18=PROVTY6_M18
	53_M18 in 2021 file	7_M18 for 2021	JTPAIN53_M18=JTPAIN7_M18
	31, 31X, or 31H in 2022 file	7, 7X, 7H for 2022	RTHLTH31= RTHLTH7 (2022 file if sample person is in 2022 data)
	42, 42X, or 42H in 2022 file	8, 8X, 8H for 2022	RTHLTH42=RTHLTH8 (2022 file)
	53, 53X, or 53H in 2022 file	9, 9X, 9H for 2022	RTHLTH53=RTHLTH9
	42_M18 in 2022 file	8_M18 for 2022	PROVTY42_M18=PROVTY8_M18
	42_M20 in 2022 file	8_M20 for 2022	ADRNK442_M20=ADRNK48_M20
	53_M23 in 2022 file	9_M23 for 2022	DENTIN53_M23=DENTIN9_M23

Type of Variable	Full-Year Consolidated PUF Variable Name Suffix	Longitudinal PUF Variable Name Suffix	Specific cases or examples
Diabetes preventive care	1853, 1953, and 2053 in 2019 file	Y0R3 for 2018 Y1R3 for 2019 Y2R3 for 2020	Example: DSEY1853=DSEYY0R3 (2019 file) DSEY1953=DSEYY1R3 (2019 file) DSEY2053=DSEYY2R3 (2019 file)
	1953, 2053, and 2153 in 2020 file	Y1R5 for 2019 Y2R5 for 2020 Y3R5 for 2021	DSEY1953=DSEYY1R5 (2020 file) DSEY2053=DSEYY2R5 (2020 file) DSEY2153=DSEYY3R5 (2020 file)
	2053, 2153, and 2253 in 2021 file	Y2R7 for 2020 Y3R7 for 2021 Y4R7 for 2022	DSEY2053=DSEYY2R7 (2021 file) DSEY2153=DSEYY3R7 (2021 file) DSEY2253=DSEYY4R7 (2021 file)
	2153, 2253, and 2353 in 2022 file	Y3R9 for 2021 Y4R9 for 2022 Y5R9 for 2023	DSEY2153=DSEYY3R9 (2022 file) DSEY2253=DSEYY4R9 (2022 file) DSEY2353=DSEYY5R9 (2022 file)
Job Change	3142 or 4253	12 for 2019 23 for 2019	All variables: CHGJ3142=CHGJ12 (2019 file) CHGJ4253=CHGJ23 (2019 file) YCHJ3142=YCHJ12 (2019 file) YCHJ4253=YCHJ23 (2019 file)
		34 for 2020 45 for 2020	CHGJ3142=CHGJ34 (2020 file) CHGJ4253=CHGJ45 (2020 file) YCHJ3142=YCHJ34 (2020 file) YCHJ4253=YCHJ45 (2020 file)
		56 for 2021 67 for 2021	CHGJ3142=CHGJ56 (2021 file) CHGJ4253=CHGJ56 (2021 file) YCHJ3142=YCHJ67 (2021 file) YCHJ4253=YCHJ67 (2021 file)
		78 for 2022 89 for 2022	CHGJ3142=CHGJ78 (2022 file) CHGJ4253=CHGJ78 (2022 file) YCHJ3142=YCHJ89 (2022 file) YCHJ4253=YCHJ89 (2022 file)
Cancer	No suffixes[4]	Y1 for 2019	Example: CALUNG=CALUNGY1 (2019 file)
		Y2 for 2020	CALUNG=CALUNGY2 (2020 file)
		Y3 for 2021	CALUNG=CALUNGY3 (2021 file)
		Y4 for 2022	CALUNG=CALUNGY4 (2022 file)

[4] To maintain a previously implemented 8-character naming convention, some variable names had the last character or two dropped in the renaming process.

Type of Variable	Full-Year Consolidated PUF Variable Name Suffix	Longitudinal PUF Variable Name Suffix	Specific cases or examples
Age of Diagnosis	No suffixes[4]	Y1 for 2019	Example: CHDAGED=CHDAGY1 (2019 file) CHOLAGED=CHOLAGY1(2019 file)
		Y2 for 2020	CHDAGED=CHDAGY2 (2020 file) CHOLAGED=CHOLAGY2(2020 file)
		Y3 for 2021	CHDAGED=CHDAGY3 (2021 file) CHOLAGED=CHOLAGY3(2021 file)
		Y4 for 2022	CHDAGED=CHDAGY4 (2022 file) CHOLAGED=CHOLAGY4(2022 file)
SDOH[5]	No suffixes[6]	5	Example: SDOHELIG=SDOHELIG5 SDAFRDHOME=SDAFRDHOME5

2.1.2 Constructed Variables for Selection of Group

The following eight variables were constructed and included on the file to facilitate the selection of appropriate cases for various analyses. Table 2 below contains descriptive statistics for these variables.

YEARIND 1=All four years, 2=2019 only, 3=2020 only, 4=2021 only, 5=2022 only, 6=2019 and 2020 only, 7=2019 and 2021 only, 8=2019 and 2022 only, 9=2020 and 2021 only, 10=2020 and 2022 only, 11= 2021 and 2022 only, 12=2019, 2020, and 2021 only, 13=2019, 2020, and 2022 only, 14=2019, 2021, and 2022 only, 15=2020, 2021, and 2022 only

ALL9RDS In scope and data collected in all nine rounds (0=no, 1=yes)

DIED Died during the four-year survey period (0=no, 1=yes)

INST Institutionalized for some time during the four-year survey period (0=no, 1=yes)

MILITARY Active duty military for some time during the four-year survey period (0=no, 1=yes)

ENTRSRVY Entered survey after beginning of panel (mainly births; also includes persons who had no initial chance of selection who moved into a MEPS sample household) (0=no, 1=yes)

LEFTUS Moved out of the country after beginning of panel (0=no, 1=yes)

OTHER Not identified in any of the above analytic groups (0=no, 1=yes)

Table 2. Frequencies and Percentage for Constructed Variables

Variable	Number of Records	Percentage of Records (N=7,080)

[5] The SDOH survey was fielded during Panel 24 Round 5 of the MEPS data collection.

YEARIND=1 (i.e., person in all four years)	5,108	91.79%
ALL9RDS=1 (yes)	4,883	87.74%
DIED=1 (yes)	297	5.34%
INST=1 (yes)	47	0.84%
MILITARY=1 (yes)	25	0.45%
ENTRSRVY=1 (yes)	248	4.46%
LEFTUS=1 (yes)	26	0.47%
OTHER=1 (yes)	52	0.93%

Following are examples of situations where these variables would be useful in selecting records for analysis:

- Analysts interested in working only with persons who were in-scope and had data for all nine rounds of the panel should subset to cases where ALL9RDS=1.
- If a researcher wanted to include persons who were in-scope and had data for all nine rounds of the panel as well as those in the survey at the beginning of the panel who subsequently died, then they would include cases where ALL9RDS=1 or (ENTRSRVY=0 and DIED=1).
- If a researcher wanted to include persons who were in-scope and had data for all nine rounds of the panel, as well as those who were in-scope across all four years but died in the fourth year of the panel, then they would include cases where ALL9RDS=1 or (DIED=1 and YEARIND=1).

2.1.3 Estimation Variables

Longitudinal Estimations for Panel 24

The file contains a weight variable (LONGWT) and variance estimation variables (VARSTR, VARPSU) that should be applied when producing national estimates for longitudinal analyses. For example, LONGWT applied to the 4,883 cases where ALL9RDS=1 produces a weighted population estimate of 303.7 million. This represents an estimate of the number of persons in the civilian noninstitutionalized population for the entire four-year period from 2019-2022. To obtain estimates of variability (such as the standard error of sample estimates or corresponding confidence intervals) for estimates based on MEPS survey data, one needs to take into account the complex sample design of MEPS by specifying the estimation variables including stratum of sample selection (VARSTR), primary sampling unit (VARPSU) and longitudinal weight (LONGWT).

This longitudinal file also contains a longitudinal SAQ weight variable (LSAQWT). This weight variable should be used to perform longitudinal analyses involving any variables from the self-administered questionnaire (SAQ) which was administered to persons age 18 and older in Rounds 2, 4, 6, and 8 of the survey. The variable SAQRDS2468 can be used to identify which persons have SAQ data for all four rounds. Similarly, the variable SAQRDS246 can be used to identify which persons have SAQ data for Rounds 2, 4, and 6; the variable SAQRDS468 can be used to identify which persons have SAQ data for Rounds 4, 6, and 8; the variable SAQRDS24 can be used to identify which persons have SAQ data for Rounds 2 and 4; the variable

SAQRDS46 can be used to identify which persons have SAQ data for Rounds 4 and 6; and the variable SAQRDS68 can be used to identify which persons have SAQ data for Rounds 6 and 8. Table 3 below provides the estimated population size (i.e., the sum of LSAQWT values) for cases with all four rounds of SAQ data (i.e., SAQRDS2468=1) and for cases with two or three rounds of SAQ data. The estimated population size for analyses based on the 1,990 cases with SAQ data for all four rounds (i.e., SAQRDS2468=1) is 176.75 million.

Table 3. Number of Respondents and Estimated Population Size for SAQ Analyses

SAQ Variable	Value	Description	Number of Respondents (Unweighted)	Estimated Population Size (Weighted by LSAQWT)
Total	Total	All SAQ respondents	5,565	259,078,778
SAQRDS2468	0	Persons with less than four rounds of SAQ data	3,575	82,327,093
SAQRDS2468	1	Persons with all four rounds of SAQ data	1,990	176,751,685
SAQRDS246	0	Persons without Rounds 2, 4, and 6 of SAQ data	3,131	81,829,266
SAQRDS246	1	Persons with Rounds 2, 4, and 6 of SAQ data	2,434	177,249,512
SAQRDS468	0	Persons without Rounds 4, 6, and 8 of SAQ data	3,481	70,744,153
SAQRDS468	1	Persons with Rounds 4, 6, and 8 of SAQ data	2,084	188,334,625
SAQRDS24	0	Persons without Rounds 2 and 4 of SAQ data	2,642	60,717,672
SAQRDS24	1	Persons with Rounds 2 and 4 of SAQ data	2,923	198,361,105
SAQRDS46	0	Persons without Rounds 4 and 6 of SAQ data	2,992	70,196,425
SAQRDS46	1	Persons with Rounds 4 and 6 of SAQ data	2,573	188,882,352
SAQRDS68	0	Persons without Rounds 6 and 8 of SAQ data	3,263	41,166,784
SAQRDS68	1	Persons with Rounds 6 and 8 of SAQ data	2,302	217.911,994

Pooled Estimations

When analyzing subpopulations and/or low-prevalence events, it may be necessary to pool together data from multiple MEPS-HC panels to accumulate a large enough sample size for producing reliable estimates. Panel 24 is the second panel to include four years of data, so this four-year file should only be combined with the Panel 23 4-year Longitudinal PUF (HC-236). However, the two-year Panel 24 Longitudinal PUF (HC-217) may be pooled with other two-year longitudinal data files, and the three-year Panel 24 Longitudinal PUF (HC-235) may be pooled with the three-year Panel 23 Longitudinal PUF (HC-226).

To ensure accurate variance estimation in such pooled analyses, a consistent and appropriate variance structure must be applied. MEPS longitudinal weight files for Panels 1–6 were released using panel-specific variance structures. Beginning with Panel 7, however, longitudinal files adopted a common variance structure. This common structure was subsequently revised starting with Panel 24.

To ensure correct variance estimation when pooling longitudinal files, the guidance below should be followed:

1. **Pooling within Panels 7–23 or within Panels 24 and beyond:**
   Simply use the variance strata and PSU variables (**VARSTR**, **VARPSU**)[6] provided on the longitudinal files.

2. **Pooling that involves either:**
   a) Any panel from Panels 1–6, or
   b) Any earlier panel in combination with Panels 24 and beyond:
   Use the variance structure from the pooled linkage public use file **HC-036**, which contains the appropriate consistent variance structure for such combinations.

The **HC-036** file is updated annually to include the correct variance structures through the most recent year. Additional information, including a summary chart outlining the appropriate variance structures for various pooling scenarios, can be found in the public use documentation for **HC-036** (see Page C-1 for the chart).

[6] Note that variable names for strata and PSU are VARSTR and VARPSU, respectively, in longitudinal files for Panel 9 and beyond. These variables were named differently in the longitudinal files for Panel 7 (VARSTRP7, VARPSUP7) and Panel 8 (VARSTRP8, VARPSUP8) and need to be standardized when pooling with subsequent panels.