MEPS HC-183:
Panel 19 Longitudinal Data File

September 2017

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

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.

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

B.1 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 health care. Estimates can be produced for individuals, families, and selected population subgroups. The panel design of the survey, which includes 5 Rounds of interviews covering 2 full calendar years, provides 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 and reflects an oversample of blacks and Hispanics. 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. MEPS further oversamples additional policy relevant sub-groups such as low income households. The linkage of the MEPS to the previous year's NHIS provides additional data for longitudinal analytic purposes.

B.2 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 cannot accurately provide. This part of the MEPS is called the Medical Provider Component (MPC) and information is collected on dates of visit, 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 date filled and sources and 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.

B.3   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. 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 summary reports, micro data files, and tables via the MEPS website: https://meps.ahrq.gov. Selected data can be analyzed through MEPSnet, an on-line interactive tool designed to give data users the capability to statistically analyze MEPS data in a menu-driven environment.

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 Ln, Rockville, MD 20857 (301-427-1406).

<center>**C. Technical and Programming Information**</center>

C.1 General Information

This documentation describes the Panel 19 longitudinal data file from the Medical Expenditure Panel Survey Household Component (MEPS-HC). Released as an ASCII file (with related SAS, STATA, and SPSS programming statements and data use information) and a SAS transport dataset, this public use file provides information collected on a nationally representative sample of the civilian noninstitutionalized population of the United States for the two-year period 2014-2015. The file contains 3,489 variables and has a logical record length of 10,139 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 1-5 of MEPS Panel 19 and can be used to analyze changes over a two-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 2014 and 2015 Full-Year Consolidated Files (HC-171 and HC-181, 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 19 data file is provided in a separate file (H183CB.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: https://meps.ahrq.gov.

C.2 Data File Information

This public use file contains records for 15,898 persons in Panel 19 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 two-year period. Only persons with positive person-level weights (PERWT14F or PERWT15F) are included in the longitudinal PUF data. Data are available for all five rounds for 92.1% of the cases (14,643). The remaining 7.9% (1,255 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 two-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 two-year period.

Each MEPS panel can be linked back to the previous year's National Health Interview Survey public use data files. For information on obtaining MEPS/NHIS link files please see https://meps.ahrq.gov/mepsweb/data_stats/more_info_download_data_files.jsp.

C.2.1 Variables

C.2.1.1 Variables from Annual Full-year Consolidated Files

Most variables on this file were obtained from the MEPS 2014 and 2015 Full-Year Consolidated Files (HC-171 and HC-181, respectively). However, names for time dependent variables from these 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 "14" and "15" are renamed with a suffix of "Y1" and "Y2", 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 2014 full-year file and "3", "4", or "5" for variables originating from Rounds 3-5 on the 2015 full-year file).[2] It is necessary to use this crosswalk in conjunction with documentation for the 2014 and 2015 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 Longitudinal File

Type of Variable	Full-Year Consolidated File Variable Name Suffix	Longitudinal File Variable Name Suffix	Specific cases or examples
Constant (i.e., not round or year specific)	No suffixes	No suffixes	All variables: DOBMM=DOBMM DOBYY=DOBYY DUID=DUID PID=PID DUPERSID=DUPERSID EDRECODE=EDRECODE EDUYRDG=EDUYRDG EDUCYR=EDUCYR HIDEG=HIDEG HISPANX=HISPANX HISPNCAT=HISPNCAT INTVLANG=INTVLANG RACEAX=RACEAX RACEBX=RACEBX RACEWX=RACEWX RACEV1X=RACEV1X

---

[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.

[2] While round 3 values were obtained for most observations from the 2015 Full Year Consolidated File, they were obtained from the 2014 Full Year Consolidated File for sample persons where YEARIND=2 (i.e., in 2014 only).

			RACEV2X=RACEV2X RACETHX=RACETHX SEX=SEX VARPSU=VARPSU VARSTR=VARSTR PANEL=PANEL BORNUSA=BORNUSA HWELLSPE=HWELLSPE LANGSPK=LANGSPK OTHLANG=OTHLANG YRSINUS=YRSINUS
Annual, family related variables	YR	Y1 or YR1 Y2 or YR2	All variables: FAMIDYR=FAMIDYR1 (2014file) FAMRFPYR=FAMRFPY1 (2014file) FAMSZEYR=FAMSZYR1 (2014 file) FAMIDYR=FAMIDYR2 (2015 file) FAMRFPYR=FAMRFPY2 (2015 file) FAMSZEYR=FAMSZYR2 (2015 file)
Annual, CPS family identifiers	No suffix	Y1 Y2	All variables: CPSFAMID= CPSFAMY1 (2014) CPSFAMID= CPSFAMY2 (2015)
Annual, health insurance eligibility units	No suffix	Y1 Y2	All variables: HIEUIDX=HIEUIDY1 (2014) HIEUIDX=HIEUIDY2 (2015)
Annual, inscope variables	No suffixes	YR1 YR2	All variables: INSCOPE=INSCPYR1 (2014 file) INSCOPE=INSCPYR2 (2015 file)
12/31 status variables	1231 in 2014 file 1231 in 2015 file	Y1 Y2	All variables: FAMS1231=FAMSY1 (2014 file) FCRP1231=FCRPY1 (2014 file) FCSZ1231= FCSZY1 (2014 file) FMRS1231= FMRSY1 (2014 file) INSC1231=INSCY1 (2014 file) FAMS1231=FAMSY2 (2015 file) FCRP1231=FCRPY2 (2015 file) FCSZ1231= FCSZY2 (2015 file) FMRS1231= FMRSY2 (2015 file) INSC1231=INSCY2 (2015 file)
Annual	14, 14X, 14F, or 14C 15, 15X, 15F, or 15C	Y1, Y1X, Y1F, or Y1C Y2, Y2X, Y2F, or Y2C	Examples: TOTEXP14=TOTEXPY1 AGE14X=AGEY1X TOTEXP15=TOTEXPY2 AGE15X=AGEY2X

Variables for health insurance prior to January 1, 2014 (data collected in round 1 only)	No suffixes	No suffixes	All variables: PREVCOVR=PREVCOVR COVRMM=COVRMM COVRYY=COVRYY WASESTB=WASESTB WASMCARE=WASMCARE WASMCAID=WASMCAID WASCHAMP=WASCHAMP WASVA=WASVA WASPRIV=WASPRIV WASOTGOV=WASOTGOV WASAFDC=WASAFDC WASSSI=WASSSI WASSTAT1=WASSTAT1 WASSTAT2=WASSTAT2 WASSTAT3=WASSTAT3 WASSTAT4=WASSTAT4 WASOTHER=WASOTHER NOINSBEF=NOINSBEF NOINSTM=NOINSTM NOINUNIT=NOINUNIT MORECOVR=MORECOVR INSENDMM=INSENDMM INSENDYY=INSENDYY
Annual	No suffixes[3]	Y1 Y2	All variables: KEYNESS=KEYNESY1 (2014 file) SAQELIG=SAQELIY1 (2014 file) EVRWRK=EVRWRKY1 (2014 file) EVRETIRE=EVRETIY1 (2014 file) AGELAST=AGELSTY1 (2014 file) KEYNESS=KEYNESY2 (2015 file) SAQELIG=SAQELIY2 (2015 file) EVRWRK=EVRWRKY2 (2015 file) EVRETIRE=EVRETIY2 (2015 file) AGELAST=AGELSTY2 (2015 file)
Monthly	2-character month + 14 2-character month + 15	2-character month + Y1 2-character month + Y2	Example: PRIJA14=PRIJAY1 (2014 file) PRIJA15=PRIJAY2 (2015 file)
Round Specific	31, 31X, or 31H in 2014 42, 42X, or 42H	1, 1X, 1H for 2014 2, 2X, 2H for	Example: RTHLTH31 = RTHLTH1 (2014 file) RTHLTH42 =RTHLTH2 (2014 file) RTHLTH53 =RTHLTH3 (2014 file if

---

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

	in 2014 53 ,53X, 53H in 2014 31, 31X, or 31H in 2015 42, 42X, or 42H in 2015 53, 53X, or 53H in 2015	2014 3, 3X, 3Hfor 2014 3, 3X, 3H for 2015 4, 4X, 4H for 2015 5, 5X, 5H for 2015	YEARIND=2) RTHLTH31 = RTHLTH3 (2015 file if YEARIND=1 or 3) RTHLTH42 =RTHLTH4 (2015 file) RTHLTH53 =RTHLTH5 (2015 file)
Diabetes preventive care	1353, 1453, and 1553 in 2014 file 1453, 1553, and 1653 in 2015 file	Y0R3 for 2013 Y1R3 for 2014 Y2R3 for 2015 Y1R5 for 2014 Y2R5 for 2015 Y3R5 for 2016	Example: DSEB1353=DSEBY0R3 (2014 file) DSEY1353=DSEYY0R3 (2014 file) DSEY1453=DSEYY1R3 (2014 file) DSEY1553=DSEYY2R3 (2014 file) DSEB1453=DSEBY1R5 (2015 file) DSEY1453=DSEYY1R5 (2015 file) DSEY1553=DSEYY2R5 (2015 file) DSEY1653=DSEYY3R5 (2015 file)
Job Change	3142 4253	12 for 2014 23 for 2014 34 for 2015 45 for 2015	All cases: CHGJ3142=CHGJ12(2014 file) CHGJ4253=CHGJ23(2014 file) YCHJ3142=YCHJ12(2014 file) YCHJ4253=YCHJ23(2014 file) CHGJ3142=CHGJ34 (2015 file) CHGJ4253=CHGJ45 (2015 file) YCHJ3142=YCHJ34 (2015 file) YCHJ4253=YCHJ45 (2015 file)
Cancer/ Cancer in remission[4]	No suffixes[5]	Y1 for 2014 Y2 for 2015	Example: CALUNG=CALUNGY1(2014 file) CALUNG=CALUNGY2 (2015 file)
Age of Diagnosis	No suffixes[5]	Y1 for 2014 Y2 for 2015	Example: CHDAGED=CHDAGY1 (2014 file) CHDAGED=CHDAGY2 (2015 file) CHOLAGED=CHOLAGY1(2014 file) CHOLAGED=CHOLAGY2(2015 file)

---

[4] Starting in 2010, variables were added to indicate whether each reported cancer was in remission.

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

C.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=both years, 2=in 2014 only, and 3=in 2015 only
ALL5RDS   Inscope and data collected in all 5 rounds (0=no, 1=yes)
DIED   Died during the two-year survey period (0=no, 1=yes)
INST   Institutionalized for some time during the two-year survey period (0=no, 1=yes)
MILITARY   Active duty military for some time during the two-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=15,898)
YEARIND=1 (i.e., person in both years)	15,545	97.8
ALL5RDS=1 (yes)	14,643	92.1
DIED=1 (yes)	180	1.1
INST=1 (yes)	60	0.4
MILITARY=1 (yes)	33	0.2
ENTRSRVY=1 (yes)	878	5.5
LEFTUS=1 (yes)	55	0.3
OTHER=1 (yes)	71	0.4

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 five rounds of the panel should subset to cases where ALL5RDS=1.
- If a researcher wanted to include persons who were in-scope and had data for all five 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 ALL5RDS=1 or (ENTRSRVY=0 and DIED=1).
- If a researcher wanted to include persons who were in-scope and had data for all five rounds of the panel as well as those who died in the second year of the panel, then they would include cases where ALL5RDS=1 or (DIED=1 and YEARIND=1).

C.2.1.3 Estimation Variables

Longitudinal Estimations for Panel 19
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 14,643 cases where ALL5RDS=1 produces a weighted population estimate of 298.3 million. This represents an estimate of the number of persons in the civilian noninstitutionalized population for the entire two-year period from 2014-2015. 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 both rounds 2 and 4 of the survey. The variable SAQRDS24 can be used to identify which persons have SAQ data for both versus only one of the two rounds. The estimated population size (i.e. sum of LSAQWT values) for analyses based on the 9,322 cases with SAQ data for both rounds (i.e., SAQRDS24=1) is 227,608,700.

Pooled Estimations
When analyzing subpopulations and/or low prevalence events, it may be desirable to pool together more than one panel of MEPS-HC data to yield sample sizes large enough to generate reliable estimates. If only data from Panels 7 and beyond are being pooled, then simply use the strata and PSU variables (VARSTR, VARPSU)[6] provided on the longitudinal files for pooled estimation. However, because Panels 1-6 MEPS longitudinal weight files were released with panel-specific variance structures, it is necessary to obtain the set of appropriate variance estimation variables from the HC-036 Pooled Estimation File when pooling involves these panels.

---

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.