MEPS HC-86:
Panel 8 Longitudinal Data File

October 2012

Agency for Healthcare Research and Quality
Center for Financing, Access, and Cost Trends
540 Gaither Road
Rockville, MD 20850
(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 five 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 non-institutionalized population and reflects an oversample of blacks and Hispanics.

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 web site: http://www.meps.ahrq.gov/mepsweb. 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, 540 Gaither Road, Rockville, MD 20850 (301-427-1406).

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

C.1   General Information

For MEPS Panels 1-8, the original longitudinal weight files that were released contained a limited number of variables that could be merged with data from two consecutive full-year consolidated files to create a longitudinal file for analysis. Beginning with Panel 9, AHRQ began to create the longitudinal weight files with more complete and analytically useful panel-specific files that contain the variables from the consolidated full-year files. AHRQ has revisited Panels 1-8 to replace the original longitudinal weight files with expanded versions, starting with Panel 8 and moving backwards.

This documentation describes the Panel 8 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 non-institutionalized population of the United States for the two-year period 2003-04. The file contains 3,107 variables and has a logical record length of 8,399 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 8 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 2003 and 2004 Full-Year Consolidated Files (HC-79 and HC-89, 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 8 data file is provided in a separate file (H86CB.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 Web site: http://www.meps.ahrq.gov/mepsweb .

C.2   Data File Information

This public use file contains records for 16,256 persons in Panel 8 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. Data are available for all five rounds for 91.1% of the cases (14,801). The remaining 8.9% (1,455 persons) do not have data for one or more rounds but participated in the survey for their full period of eligibility. These persons include 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 did not participate in the survey for the entire period they were in-scope are not included in this file. The analytic weight variable (LONGWT) has been adjusted for nonresponse/attrition and should be used to produce national estimates for the two-year period. The codebook provides both weighted and unweighted frequencies for each variable in the data file.

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: http://www.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 2003 and 2004 Full-Year Consolidated Files (HC-79 and HC-89, 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 "03" and "04" 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 2003 full-year file and "3", "4", or "5" for variables originating from Rounds 3-5 on the 2004 full-year file).[2] It is necessary to use this crosswalk in conjunction with documentation for the 2003 and 2004 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	Examples
Constant (i.e., not round or year specific)	No suffixes	No suffixes	DOBMM=DOBMM DOBYY=DOBYY DUID=DUID PID=PID DUPERSID=DUPERSID
Annual, family related variables	YR	Y1 or YR1 Y2 or YR2	FAMIDYR=FAMIDYR1 (2003 file) FAMRFPYR=FAMRFPY1 (2003file) FAMSZEYR=FAMSZYR1 (2003 file) FAMIDYR=FAMIDYR2 (2004 file) FAMRFPYR=FAMRFPY2 (2004 file) FAMSZEYR=FAMSZYR2 (2004 file)
Annual,	No suffix	Y1	CPSFAMID= CPSFAMY1 (2003)

---

[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 2004 Full Year Consolidated File, they were obtained from the 2003 Full Year Consolidated File for sample persons where YEARIND=2 (i.e., in 2003 only).

CPS family identifiers		Y2	CPSFAMID= CPSFAMY2 (2004)
Annual, health insurance eligibility units	No suffix	Y1 Y2	HIEUIDX=HIEUIDY1 (2003) HIEUIDX=HIEUIDY2 (2004)
Annual, inscope variables	No suffixes	YR1 YR2	INSCOPE=INSCPYR1 (2003 file) INSCOPE=INSCPYR2 (2004 file)
12/31 status variables	1231 in 2003 file 1231 in 2004 file	Y1 Y2	FAMS1231=FAMSY1 (2003 file) FCRP1231=FCRPY1 (2003 file) FCSZ1231= FCSZY1 (2003 file) FMRS1231= FMRSY1 (2003 file) INSC1231=INSCY1 (2003 file) FAMS1231=FAMSY2 (2004 file) FCRP1231=FCRPY2 (2004 file) FCSZ1231= FCSZY2 (2004 file) FMRS1231= FMRSY2 (2004 file) INSC1231=INSCY2 (2004 file)
Annual	03, 03X, 03F, or 03C 04, 04X, 0402F, or 04C	Y1, Y1X, Y1F, or Y1C Y2, Y2X, Y2F, or Y2C	TOTEXP03=TOTEXPY1 (2003 file) AGE03X=AGEY1X TOTEXP04=TOTEXPY2 (2004 file) AGE04X=AGEY2X
Variables for health insurance prior to January 1, 2003 (data collected in round 1 only)	No suffixes	No suffixes	PREVCOVR=PREVCOVR COVRMM=COVRMM COVRYY=COVRYY WASESTB=WASESTB WASMCARE=WASMCARE WASMCAID=WASMCAID WASCHAMP=WASCHAMP WASVA=WASVA WASPRIV=WASPRIV WASOTGOV=WASOTGOV WASAFDC=WASAFDC WASSSI=WASSSI WASOTHER=WASOTHER NOINSBEF=NOINSBEF NOINSTM=NOINSTM NOINUNIT=NOINUNIT MORECOVR=MORECOVR INSENDMM=INSENDMM INSENDYY=INSENDYY

Annual	No suffixes[3]	Y1	KEYNESS=KEYNESY1 (2003 file)
			EVRWRK=EVRWRKY1 (2003 file)
			EVRETIRE=EVRETIY1 (2003 file)
		Y2	KEYNESS=KEYNESY2 (2004 file)
			EVRWRK=EVRWRKY2 (2004 file)
			EVRETIRE=EVRETIY2 (2004 file)
Monthly	2-character month + 03	2-character month + Y1	PRIJA03=PRIJAY1 (2003 file)
	2-character month + 04	2-character month + Y2	PRIJA04=PRIJAY2 (2004 file)
Round Specific	31 or 31X in 2003	1 or 1X for 2003	RTHLTH31 = RTHLTH1 (2003 file)
	42 or 42X in 2003	2 or 2X for 2003	RTHLTH42 =RTHLTH2 (2003 file)
	53 or 53X in 2003	3 or 3X for 2003	RTHLTH53 =RTHLTH3 (2003 file if YEARIND=2)
	31 or 31X in 2004	3 or 3X for 2004	RTHLTH31 = RTHLTH3 (2004 file if YEARIND=1 or 3)
	42 or 42X in 2004	4 or 4X for 2004	RTHLTH42 =RTHLTH4 (2004 file)
	53 or 53X in 2004	5 or 5X for 2004	RTHLTH53 =RTHLTH5 (2004 file)
Diabetes preventive care	0253, 0353, and 0453 in 2003 file	Y0R3 for 2002	DSEB0253=DSEBY0R3 (2003 file)
		Y1R3 for 2003	DSEY0253=DSEYY0R3 (2003 file)
		Y2R3 for 2004	DSEY0353=DSEYY1R3 (2003 file)
			DSEY0453=DSEYY2R3 (2003 file)
	0353, 0453, and 0553 in 2004 file	Y1R5 for 2003	DSEB0353=DSEBY1R5 (2004 file)
		Y2R5 for 2004	DSEY0353=DSEYY1R5 (2004 file)
		Y3R5 for 2005	DSEY0453=DSEYY2R5 (2004 file)
			DSEY0553=DSEYY3R5 (2004 file)
Job Change	3142	12 for 2003	CHGJ3142=CHGJ12(2003 file)
	4253	23 for 2003	CHGJ4253=CHGJ23(2003 file)
			YCHJ3142=YCHJ12(2003 file)
			YCHJ4253=YCHJ23(2003 file)
		34 for 2004	CHGJ3142=CHGJ34 (2004 file)
		45 for 2004	CHGJ4253=CHGJ45 (2004 file)
			YCHJ3142=YCHJ34 (2004 file)
			YCHJ4253=YCHJ45 (2004 file)

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.

---

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

YEARIND   1=both years, 2=in 2003 only, and 3=in 2004 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=16,256)
YEARIND=1 (i.e., person in both years)	15,774	97.0
ALL5RDS=1 (yes)	14,801	91.1
DIED=1 (yes)	212	1.3
INST=1 (yes)	84	0.5
MILITARY=1 (yes)	37	0.2
ENTRSRVY=1 (yes)	970	6.0
LEFTUS=1 (yes)	84	0.5
OTHER=1 (yes)	89	0.6

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 8
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 14,801 cases where ALL5RDS=1 produces a weighted

population estimate of 272.1 million. This represents an estimate of the number of persons in the civilian non-institutionalized population for the entire two-year period from 2003-04. 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).

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) 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.