

**MEPS HC-067H:  
2002 Home Health Visits**

**September 2004**

**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 .....	A-1
B. Background .....	B-1
1.0 Household Component .....	B-1
2.0 Medical Provider Component.....	B-2
3.0 Insurance Component .....	B-2
4.0 Survey Management.....	B-3
C. Technical And Programming Information .....	C-1
1.0 General Information .....	C-1
2.0 Data File Information .....	C-2
2.1 Using MEPS Data for Trend and Longitudinal Analysis.....	C-4
2.2 Codebook Structure.....	C-5
2.3 Reserved Codes .....	C-5
2.4 Codebook Format .....	C-6
2.5 Variable Source and Naming Conventions .....	C-6
2.5.1 Variable-Source Crosswalk .....	C-6
2.5.2 Expenditure and Source of Payment Variables .....	C-7
2.6 File Contents .....	C-8
2.6.1 Survey Administration Variables .....	C-8
2.6.1.1 Person Identifiers (DUID, PID, DUPERSID).....	C-8
2.6.1.2 Record Identifier (EVNTIDX) .....	C-8
2.6.1.3 Round Indicator (EVENTRN).....	C-8
2.6.2 Home Health Event Variables .....	C-8
2.6.2.1 Date of Event (HHDATEYR, HHDATEMM).....	C-8
2.6.2.2 Characteristics of Event (MPCELIG-OTHWCOS) .....	C-9
2.6.2.3 Treatments, Therapies, and Services (HOSPITAL-OTHSVCOS) .....	C-9
2.6.2.4 Frequency of Event (FREQCY-HHDAYS) .....	C-10
2.6.3 Condition, Procedure, and Clinical Classification Codes.....	C-10
2.6.4 Expenditure Data .....	C-10
2.6.4.1 Definition of Expenditures .....	C-10
2.6.4.2 Data Editing and Imputation Methodologies of Expenditure Variables .....	C-11
2.6.4.2.1 General Data Editing Methodology .....	C-11
2.6.4.2.2 General Hot-Deck Imputation .....	C-12
2.6.4.2.3 Home Health Data Editing and Imputation .....	C-12
2.6.4.3 Imputation Flag Variable (IMPFLAG) .....	C-13
2.6.4.4 Zero Expenditures .....	C-13
2.6.4.5 Sources of Payment .....	C-14
2.6.4.6 Home Health Expenditure Variables (HHSF02X - HHXP02X).....	C-15
2.6.4.7 Rounding .....	C-15

3.0	Sample Weight (PERWT02F)	C-16
3.1	Overview	C-16
3.2	Details on Person Weight Construction	C-16
3.2.1	MEPS Panel 6 Weight	C-17
3.2.2	MEPS Panel 7 Weight	C-17
3.2.3	The Final Weight for 2002	C-17
3.2.4	Coverage	C-18
4.0	Strategies for Estimation	C-18
4.1	Variables with Missing Values	C-18
4.2	Basic Estimates of Utilization, Expenditures, and Sources of Payment	C-19
4.3	Estimates of the Number of Persons with Home Health Events Due to Hospitalization	C-20
4.4	Person-Based Ratio Estimates	C-21
4.4.1	Person-Based Ratio Estimates Relative to Persons with Home Health Events by Paid Independent Providers	C-21
4.4.2	Person-Based Ratio Estimates Relative to the Entire Population	C-21
4.5	Sampling Weights for Merging Previous Releases of MEPS Household Data with this Event File	C-22
4.6	Variance Estimation (VARPSU, VARSTR)	C-22
5.0	Merging/Linking MEPS Data Files	C-23
5.1	Linking a 2002 Person-Level File to the 2002 Home Health Event File	C-23
5.2	Linking the 2002 Home Health Event File to the 2002 Medical Conditions File	C-24
5.2.1	Limitations/Caveats of CLNK (the 2002 Medical Conditions Link File)	C-24
	References	C-25
D.	Variable-Source Crosswalk	D-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.

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**

The Medical Expenditure Panel Survey (MEPS) provides nationally representative estimates of health care use, expenditures, sources of payment, and insurance coverage for the U.S. civilian noninstitutionalized population. MEPS is cosponsored by the Agency for Healthcare Research and Quality (AHRQ) and the National Center for Health Statistics (NCHS).

MEPS is a family of three surveys. The Household Component (HC) is the core survey and forms the basis for the Medical Provider Component (MPC) and part of the Insurance Component (IC). Together these surveys yield comprehensive data that provide national estimates of the level and distribution of health care use and expenditures, support health services research, and can be used to assess health care policy implications.

MEPS is the third in a series of national probability surveys conducted by AHRQ on the financing and use of medical care in the United States. The National Medical Care Expenditure Survey (NMCES, also known as NMES-1) was conducted in 1977 and the National Medical Expenditure Survey (NMES-2) in 1987. Since 1996, MEPS continues this series with design enhancements and efficiencies that provide a more current data resource to capture the changing dynamics of the health care delivery and insurance systems.

The design efficiencies incorporated into MEPS are in accordance with the Department of Health and Human Services (DHHS) Survey Integration Plan of June 1995, which focused on consolidating DHHS surveys, achieving cost efficiencies, reducing respondent burden, and enhancing analytical capacities. To advance these goals, MEPS includes linkage with the National Health Interview Survey (NHIS) - a survey conducted by NCHS from which the sample for the MEPS HC is drawn - and enhanced longitudinal data collection for core survey components. The MEPS HC augments NHIS by selecting a sample of NHIS respondents, collecting additional data on their health care expenditures, and linking these data with additional information collected from the respondents' medical providers, employers, and insurance providers.

### **1.0 Household Component**

The MEPS HC, a nationally representative survey of the U.S. civilian noninstitutionalized population, collects medical expenditure data at both the person and household levels. The HC collects detailed data on demographic characteristics, health conditions, health status, use of medical care services, charges and payments, access to care, satisfaction with care, health insurance coverage, income, and employment.

The HC uses an overlapping panel design in which data are collected through a preliminary contact followed by a series of five rounds of interviews over a 2 ½-year period. Using computer-assisted personal interviewing (CAPI) technology, data on medical expenditures and use for two calendar years are collected from each household. This series of data collection rounds is launched each subsequent year on a new sample of households to provide overlapping panels of survey data and, when combined with other ongoing panels, will provide continuous and current estimates of health care expenditures.

The sampling frame for the MEPS HC is drawn from respondents to NHIS. NHIS provides a nationally representative sample of the U.S. civilian noninstitutionalized population, with oversampling of Hispanics and blacks.

## **2.0 Medical Provider Component**

The MEPS MPC supplements and/or replaces information on medical care events reported in the MEPS HC by contacting medical providers and pharmacies identified by household respondents. The MPC sample includes all home health agencies and pharmacies reported by HC respondents. Office-based physicians, hospitals, and hospital physicians are also included in the MPC but may be subsampled at various rates, depending on burden and resources, in certain years.

Data are collected on medical and financial characteristics of medical and pharmacy events reported by HC respondents. The MPC is conducted through telephone interviews and record abstraction.

## **3.0 Insurance Component**

The MEPS IC collects data on health insurance plans obtained through private and public-sector employers. Data obtained in the IC include the number and types of private insurance plans offered, benefits associated with these plans, premiums, contributions by employers and employees, eligibility requirements, and employer characteristics.

Establishments participating in the MEPS IC are selected through three sampling frames:

- A list of employers or other insurance providers identified by MEPS HC respondents who report having private health insurance at the Round 1 interview.
- A Bureau of the Census list frame of private sector business establishments.
- The Census of Governments from Bureau of the Census.

To provide an integrated picture of health insurance, data collected from the first sampling frame (employers and insurance providers identified by MEPS HC respondents) are linked back to data provided by those respondents. Data from the two Census Bureau sampling frames are used to produce annual national and state estimates of the supply and cost of private health insurance available to American workers and to evaluate policy issues pertaining to health insurance. National estimates of employer contributions to group insurance from the MEPS IC are used in the computation of Gross Domestic Product (GDP) by the Bureau of Economic Analysis.

The MEPS IC is an annual survey. Data are collected from the selected organizations through a prescreening telephone interview, a mailed questionnaire, and a telephone follow-up for nonrespondents.

#### **4.0 Survey Management**

MEPS data are collected under the authority of the Public Health Service Act. They are edited and published in accordance with the confidentiality provisions of this act and the Privacy Act. 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, microdata files and compendiums of tables. Data are released through MEPSnet, an online interactive tool developed to give users the ability to statistically analyze MEPS data in real time. Summary reports and compendiums of tables are released as printed documents and electronic files. Microdata files are released as electronic files.

Selected printed documents are available through the AHRQ Publications Clearinghouse. Write or call:

AHRQ Publications Clearinghouse  
Attn: (publication number)  
P.O. Box 8547  
Silver Spring, MD 20907  
800-358-9295  
410-381-3150 (callers outside the United States only)  
888-586-6340 (toll-free TDD service; hearing impaired only)

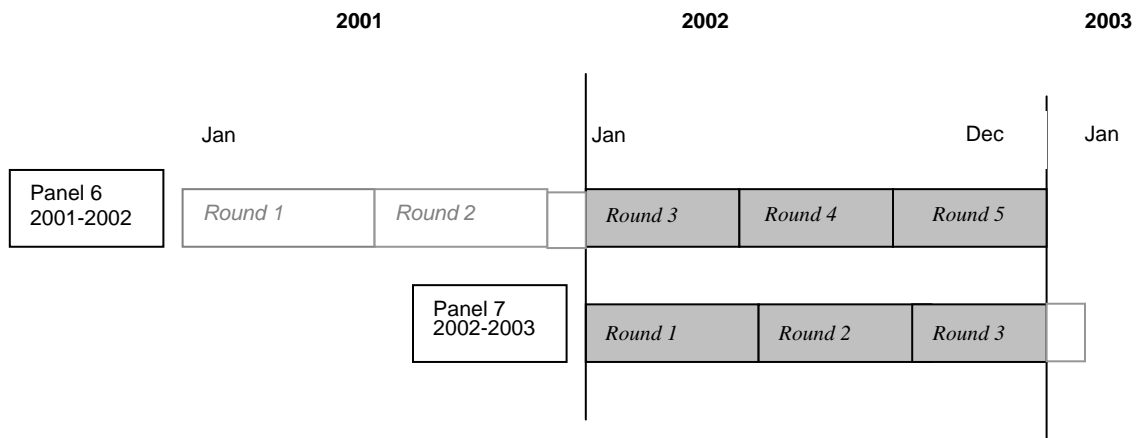
Be sure to specify the AHRQ number of the document you are requesting.

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).

## C. Technical and Programming Information

### 1.0 General Information

This documentation describes one in a series of public use event files from the 2002 Medical Expenditure Panel Survey (MEPS) Household Component (HC) and Medical Provider Component (MPC). Released as an ASCII data file (with related SAS and SPSS programming statements) and a SAS transport file, the 2002 Home Health public use file provides detailed information on home health events for a nationally representative sample of the civilian noninstitutionalized population of the United States. Data from the Home Health event file can be used to make estimates of home health event utilization and expenditures for calendar year 2002. As illustrated below, this file consists of MEPS survey data obtained in the 2002 portion of Round 3, and Rounds 4 and 5 for Panel 6, as well as Rounds 1, 2, and the 2002 portion of Round 3 for Panel 7 (i.e., the rounds for the MEPS panels covering calendar year 2002).



Counts of home health utilization are based entirely on household reports. Agency home health providers were sampled into the MEPS MPC (see Section B. 2.0). Only those providers for whom the respondent signed a permission form were included in MPC. Information from MPC was used to supplement expenditure and payment data reported by the household, and does not affect use estimates.

Data from this event file can be merged with other 2002 MEPS HC data files for the purposes of appending person-level data such as demographic characteristics or health insurance coverage to each Home Health record.

This file can also be used to construct summary variables for expenditures, sources of payment, and related aspects of home health events for calendar year 2002. Aggregate annual person-level information on the use of home health providers and other health services use is provided on the 2002 Population Characteristics File, where each record represents a MEPS sampled person.



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

- Data File Information
- Sample Weight
- Strategies for Estimation
- Merging/Linking MEPS Data Files
- References
- Variable - Source Crosswalk

For more information on MEPS HC survey design see S. Cohen, 1997; J. Cohen, 1997; and S. Cohen, 1996. For information on the MEPS MPC design, see S. Cohen, 1999. A copy of the survey instruments used to collect the information on this file is available on the MEPS web site at the following address: <<http://www.meps.ahrq.gov>>.

## **2.0 Data File Information**

The 2002 Home Health public use data set consists of one event-level data file. The file contains characteristics associated with the home health event and imputed expenditure data. For users wanting to impute expenditures, pre-imputed data are available through the Center for Financing, Access and Cost Trends (CFACT) Data Center. Please visit the CFACT Data Center web site for details: <<http://www.meps.ahrq.gov/>>.

The data user/analyst is forewarned that the imputation of expenditures will necessitate a sizable commitment of resources: financial, staff, and time.

The Home Health use data set contains characteristics associated with the home health event and imputed expenditure data. Each record represents a household-reported home health event. A home health event is a MONTH of similar services provided by the same PROVIDER -- a month of home health services from a single provider entity (i.e., paid independent, informal or agency). For example, if a person received four events from a nurse, ten events from a homemaker and four events from a physical therapist each from the same provider every month for three months, then there will be three event records (NOT 54 records) on the file, one for each month. Data were collected in this manner because agencies, hospitals, and nursing homes provide expenditure data in this manner. In order to be consistent with the definition of what is considered a home health event on this file, this same definition (i.e., a month of similar services) was applied to all types of home health providers.

This public use data set contains 5,494 home health records; of the records, 5,342 are associated with persons having a positive person-level weight (PERWT02F). It includes all records related to home health events for all household survey respondents who resided in eligible responding households and reported at least one home health event.

Each record represents one household-reported home health event that occurred during calendar year 2002. Some household respondents may have multiple events and thus will be represented in multiple records on the file. Other household respondents may have reported no events and thus will have no records on this file. These data were collected during the 2002 portion of Round 3, and Rounds 4 and 5 for Panel 6, as well as Rounds 1, 2, and the 2002 portion of Round 3 for Panel 7 of the MEPS HC. The persons represented on this file had to meet either (a) or (b):

- a) Be classified as a key in-scope person who responded for his or her entire period of 2002 eligibility (i.e., persons with a positive 2002 full-year person-level weight ( $PERWT02F > 0$ )), or
- b) Be an eligible member of a family all of whose key in-scope members have a positive person-level weight ( $PERWT02F > 0$ ). (Such a family consists of all persons with the same value for FAMIDYR.) That is, the person must a positive full-year family-level weight ( $FAMWT02F > 0$ ). Note that FAMIDYR and FAMWT02F are variables on the 2002 Population Characteristics file.

Persons with no home health events for 2002 are not included on this event-level file but are represented on the 2002 MEPS person-level file. A codebook for that data file is provided in H62CB.PDF and H62CB.ASP.

Home health providers include formal or paid, and informal or unpaid providers. Formal or paid providers include: home health agency and other independent paid providers. Informal or unpaid providers include family and friends.

For home health agencies it is important to distinguish between the provider and the home health worker. In these cases, the provider is the agency or the facility that employs the workers. The home health workers are the people who administer the care. Examples of home health care workers are the following: nurses, physical therapists, home health aides, homemakers, and hospice workers, among others. These examples are generally the types of workers associated with agencies. Paid independent providers generally include companions, nursing assistants, physicians, etc. For each record on this file, one or more types of workers can be reported. The respondent is asked to mention all of the types of home health workers who provided home health care (since records represent a month of service, there can be more than one type of worker on a single record). For example, an agency that provides two types of aides that provide home health care to the same person during a specific month is represented as one event on the file even though two workers employed at the same agency provided care. When using this file, analysts must keep in mind that a record on the file corresponds to a provider entity, not an individual or particular worker.

Expenditure data for home health agency events are collected exclusively in the MPC. Expenditure data for other paid independent home health care events are collected from the household, since these types of events are not included in the MPC. Friends, family and volunteers providing home health care to a person are considered unpaid and are not included in the MPC. No expenditure information is available for them.

Each home health record also includes the following: date the provider started seeing the respondent; type of provider; types of services provided and if this was a repeat event; whether or not care was received due to hospitalization; whether or not a person was taught how to use medical equipment; imputed sources of payment, total payment and total charge for the home health event expenditure; and a full-year person-level weight.

Data from this file can be merged with previously released 2002 MEPS HC person-level data, such as the MEPS 2002 Full-Year Population Characteristics file, using the unique person identifier, DUPERSID, to append person-level information, such as demographic or health insurance coverage, to each record. Home Health events can also be linked to the MEPS 2002 Medical Conditions File. Please see Section 5.0 or the MEPS 2002 Appendix File, HC-067I, for details on how to link MEPS data files.

## **2.1 Using MEPS Data for Trend and Longitudinal Analysis**

MEPS began in 1996 and several annual data files have been released. As more years of data are produced, MEPS will become increasingly valuable for examining health care trends. However, it is important to consider a variety of factors when examining trends over time using MEPS. Statistical significance tests should be conducted to assess the likelihood that observed trends are attributable to sampling variation. MEPS expenditure estimates are especially sensitive to sampling variation due to the underlying skewed distribution of expenditures. For example, 1 percent of the population accounts for about one-quarter of all expenditures. The extent to which observations with extremely high expenditures are captured in the MEPS sample varies from year to year (especially for smaller population subgroups), which can produce substantial shifts in estimates of means or totals that are simply an artifact of the sample(s). The length of time being analyzed should also 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 or MEPS survey methodology. Looking at changes over longer periods of time can provide a more complete picture of underlying trends. Analysts may wish to consider using techniques to smooth or stabilize trend analyses of MEPS data such as pooling time periods for comparison (e.g. 1996-97 versus 1998-99), 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, researchers should be aware of the impact of multiple comparisons on Type I error because performing numerous statistical significance tests of trends increases the likelihood of inappropriately concluding a change is statistically significant.

The records on this file can be linked to all other 2002 MEPS-HC public use data sets by the sample person identifier (DUPERSID).

## 2.2 Codebook Structure

For each variable on the Home Health event file, both weighted and unweighted frequencies are provided in the codebook (files H67HCB.PDF and H67HCB.ASP). The codebook and data file sequence list variables in the following order:

- Unique person identifier
- Unique home health event identifier
- Home health characteristic variables
- Imputed expenditure variables
- Weight and variance estimation variables

## 2.3 Reserved Codes

The following reserved code values are used:

<b>Value</b>	<b>Definition</b>
-1 INAPPLICABLE	Question was not asked due to skip pattern.
-7 REFUSED	Question was asked and respondent refused to answer question.
-8 DK	Question was asked and respondent did not know answer.
-9 NOT ASCERTAINED	Interviewer did not record the data.

Generally, values of -1, -7, -8, and -9 for non-expenditure variables have not been edited on this file. The values of -1 and -9 can be edited by the data users/analysts by following the skip patterns in the HC survey questionnaire (located on the MEPS web site: <<http://www.meps.ahrq.gov/>>).

## 2.4 Codebook Format

The codebook describes an ASCII data set (although the data are also being provided in a SAS transport file). The following codebook items are provided for each variable:

<b>Identifier</b>	<b>Description</b>
Name	Variable name (maximum of 8 characters)
Description	Variable descriptor (maximum of 40 characters)
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

## 2.5 Variable Source and Naming Conventions

In general, variable names reflect the content of the variable, with an eight-character limitation. Generally, imputed/edited variables end with an “X”.

### 2.5.1 Variable-Source Crosswalk

Variables were derived either from the HC questionnaire itself, the MPC data collection instrument, or from the CAPI. The source of each variable is identified in Section D “Variable - Source Crosswalk” in one of four ways:

- 1) Variables derived from CAPI or assigned in sampling are so indicated as “CAPI derived” or “Assigned in sampling,” respectively;
- 2) Variables which come from one or more specific questions have those questionnaire sections and question numbers indicated in the “Source” column; questionnaire sections are identified as:
  - EV – Event Roster section
  - HH – Home Health Event section
  - CP – Charge Payment section
- 3) Variables constructed from multiple questions using complex algorithms are labeled “Constructed” in the “Source” column; and
- 4) Variables that have been edited or imputed are so indicated.

## 2.5.2 Expenditure and Source of Payment Variables

The names of the expenditure and source of payment variables follow a standard convention, are seven characters in length, and end in an “X” indicating edited/imputed. Please note that imputed means that a series of logical edits, as well as an imputation process to account for missing data, have been performed on the variable.

The total sum of payments and the 12 source of payment are named in the following way:

The first two characters indicate the type of event:

IP - inpatient stay	OB - office-based visit
ER - emergency room visit	OP - outpatient visit
HH - home health event	DV - dental visit
OM - other medical equipment	RX - prescribed medicine

In the case of source of payment variables, the third and fourth characters indicate:

SF - self or family	OF - other Federal Government
MR – Medicare	SL - State/local government
MD – Medicaid	WC - Worker’s Compensation
PV - private insurance	OT - other insurance
VA – Veterans Administration	OR - other private
TR – TRICARE	OU - other public
	XP - sum of payments

In addition, the total charge variable is indicated by TC in the variable name.

The fifth and sixth characters indicate the year (02). The seventh character, “X”, indicates the variable is edited/imputed.

For example, HHSF02X is the edited/imputed amount paid by self or family for 2002 home health expenditures.

## **2.6 File Contents**

### **2.6.1 Survey Administration Variables**

#### **2.6.1.1 Person Identifiers (DUID, PID, DUPERSID)**

The dwelling unit ID (DUID) is a five-digit random number assigned after the case was sampled for MEPS. The three-digit person number (PID) uniquely identifies each person within the dwelling unit. The eight-character variable DUPERSID uniquely identifies each person represented on the file and is the combination of the variables DUID and PID. For detailed information on dwelling units and families, please refer to the documentation for the 2002 Full Year Population Characteristics file.

#### **2.6.1.2 Record Identifier (EVNTIDX)**

EVNTIDX uniquely identifies each event (i.e., each record on the home health file) and is the variable required to link home health events to data files containing details on conditions (MEPS 2002 Medical Conditions File). For details on linking see Section 5.0 or the MEPS 2002 Appendix File, HC-067I.

#### **2.6.1.3 Round Indicator (EVENTRN)**

EVENTRN indicates the round in which the home health event was reported. Please note: Rounds 3, 4, and 5 are associated with MEPS survey data collected from Panel 6. Likewise, Rounds 1, 2, and 3 are associated with data collected from Panel 7.

### **2.6.2 Home Health Event Variables**

This file contains variables describing home health events reported by household respondents in the Home Health Section of the MEPS HC survey questionnaire.

#### **2.6.2.1 Date of Event (HHDATEYR, HHDATEMM)**

The start date variables (HHDATEYR and HHDATEMM) indicate the year and month that the household respondent reported as the start date (or the first time) for this type of home health event. An artifact of the data collection for the variable HHDATEYR is that all events are reported as having started in 2002 even though a person could have started receiving that type of home health care from that provider year(s) before 2002. These variables should not be interpreted as “true” start dates.

### **2.6.2.2 Characteristics of Event (MPCELIG-OTHCWOS)**

The HC questionnaire asked the respondent to indicate whether the home health provider event(s) for each month's services were provided through an agency or an independent paid provider (SELFAGEN). The response to the SELFAGEN question dictated the skip pattern CAPI followed regarding the questions in the home health section of the HC questionnaire. The questionnaire also asked respondents if the provider was paid or whether a friend, relative, or volunteer (HHTYPE) provided the home health services. The constructed variable MPCELIG indicates whether the home health provider event was eligible for MPC data collection and the type of imputation process the event went through. MPCELIG is a more accurate variable for determining whether the event was an agency, a paid independent or an informal care event. However, SELFAGEN is a more accurate variable for determining the home health questions asked of the respondent. All respondents receiving care from an agency, hospital or nursing home were asked to identify the type of home health worker they saw (CNA-SPEECTHP) – for example, certified nursing assistant, home health aide, registered nurse, etc.

Analysts should keep in mind that these identifications by household respondents are subjective in nature, are not mutually exclusive or collectively exhaustive, and should not be used to make certain estimates. For example, a person on one type of insurance may identify an individual providing home health care services to them as a personal care attendant while an individual having a different type of insurance coverage may identify that same worker as a home care aide. Making estimates of personal care attendants or home care aides based on their identification by household respondents and treating these types of workers as mutually exclusive groups will result in inaccurate estimates. Respondents may also have indicated that they were seen by more than one home health care worker during a single event. For example, since an event is a month of services, a respondent may have reported being seen by a nurse, a physical therapist, and/or a home health aide during a single event. Respondents were also asked to identify other non-skilled, skilled, and other workers seen during that month of care (NONSKILL-OTHCWOS). However, "other specify" variables (SKILLWOS and OTHCWOS) were not reconciled with the type of health care worker variable (CNA-SPEECTHP). In addition, the type of health care worker variables (CNA-SPEECTHP) were not reconciled with MPCELIG, SELFAGEN and HHTYPE, so inconsistencies between these variables are possible.

### **2.6.2.3 Treatments, Therapies, and Services (HOSPITAL-OTHSVCOS)**

Regardless of the type of provider, all respondents were asked if the home health services received were due to a hospitalization (HOSPITAL), whether services were due to a medical condition (VSTRELCN), if the person was helped with daily activities (DAILYACT), if the person received companionship services (COMPANY), and whether or not the person received any other type of services (OTHSVCE and OTHSVCOS). Only persons receiving care from an agency were asked if they were



taught how to use medical equipment (MEDEQUIP) and whether or not they received a medical treatment (TREATMT).

#### **2.6.2.4 Frequency of Event (FREQCY-HHDAYS)**

Several variables identify the frequency and length of home health events (FREQCY-MINLONG) and whether or not the same services were received during each month (SAMESVCE). Frequency of event variables (FREQCY-TMSPDAY) were used as building blocks to construct HHDAYS. HHDAYS indicates the number of days the person received care during that event (i.e., month of care). HHDAYS has not been reconciled with DAYSPMO. Frequency variables can be combined to get a measure of the intensity of care. For example, HHDAYS can be used in conjunction with HRSLONG and TMSPDAY to form a measure of intensity of care, that is, how many hours of care were provided in one month.

#### **2.6.3 Condition, Procedure, and Clinical Classification Codes**

Information on household reported medical conditions and procedures (including condition codes, procedure codes, and clinical classification codes) associated with each home health event are NOT provided on this file. To obtain complete condition information associated with an event, the analyst must link to the 2002 Medical Conditions File. Details on how to link to the MEPS 2002 Medical Conditions File are provided in the MEPS 2002 Appendix File, HC-067I.

#### **2.6.4 Expenditure Data**

##### **2.6.4.1 Definition of Expenditures**

Expenditures on this file refer to what is paid for health care services. More specifically, expenditures in MEPS are defined as the sum of payments for care received, including out-of-pocket payments and payments made by private insurance, Medicaid, Medicare and other sources. The definition of expenditures used in MEPS differs slightly from its predecessors, 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 1990’s due to the increasingly common practice of discounting. Although measuring expenditures as the sum of payments incorporates discounts in the MEPS expenditure estimates, these estimates do not incorporate any payment not directly tied to specific medical care events, such as bonuses or retrospective payment adjustments paid by third party payers. Another general 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 charge data are provided on this file, data users/analysts should use caution when working with this data because a charge does not typically

represent actual dollars exchanged for services or the resource costs of those services, nor are they directly comparable to the expenditures defined in the 1987 NMES (for details on expenditure definitions, see Monheit et al, 1999). For details on expenditure definitions, please refer to the following, “Informing American Health Care Policy” (Monheit et al., 2000). AHRQ has developed factors to apply to the 1987 NMES expenditure data to facilitate longitudinal analysis. These factors can be accessed via the CFACT Data Center. For more information, see the Data Center section of the MEPS web site at <<http://www.meeps.ahrq.gov/>>. If examining trends in MEPS expenditures or performing longitudinal analysis on MEPS expenditures, please refer to section C, sub-section 2.1 for more information.

#### **2.6.4.2 Data Editing and Imputation Methodologies of Expenditure Variables**

The general methodology used for editing and imputing expenditure data is described below. However, please note, the MPC included home health events provided by an agency and did not include home health care provided by paid independent providers. Although the general procedures remain the same for all home health events, there were some differences in the editing and imputation methodologies applied to those events followed in the MPC and those events not followed in the MPC. Analysts should note that home health care provided by friends, family, or volunteers was assumed to be free and was not included in any imputation process. Please see below for details on the differences between these editing/imputation methodologies.

Home health expenditure data for agency, hospital, and nursing home providers were collected exclusively from the MPC (i.e., household respondents were not asked to report home health expenditures from these types of providers). The MPC contacted 100 percent of the agency, hospital, and nursing home health providers identified by household respondents. Since paid independent home health providers were not included in the MPC, all expenditure data from these providers were collected from household respondents.

##### **2.6.4.2.1 General Data Editing Methodology**

Logical edits were used to resolve internal inconsistencies and other problems in the HC and MPC survey-reported data. The edits were designed to preserve partial payment data from households and providers, and to identify actual and potential sources of payment for each household-reported event. In general, these edits accounted for outliers, co-payments 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 mis-classifications between Medicare and Medicaid and between Medicare HMO’s and private HMO’s as payment sources. These edits produced a complete vector of expenditures for some events, and provided the starting point for imputing missing expenditures in the remaining events.

#### **2.6.4.2.2 General Hot-Deck Imputation**

A weighted sequential hot-deck procedure was used to impute for missing expenditures, as well as total charge. This procedure uses survey data from respondents to replace missing data, while taking into account the respondents' weighted distribution in the imputation process. Classification variables vary by event type in the hot-deck imputations, but total charge and insurance coverage are key variables in all of the imputations. Separate imputations were performed for nine categories of medical provider care: inpatient hospital stays, outpatient hospital department visits, emergency room visits, visits to physicians, visits to non-physician providers, dental services, home health care by certified providers, home health care by paid independents, and other medical expenses. Within each file, flat fee and simple events were imputed separately. After the imputations were finished, visits to physician and non-physician providers were combined into a single medical provider file. The two categories of home care also were combined into a single home health file.

#### **2.6.4.2.3 Home Health Data Editing and Imputation**

Expenditures for home health events were developed in a sequence of logical edits and imputations. (Analysts should note that home health care provided by friends, family, or volunteers was assumed not to have associated expenditures and was not included in any imputation process. All expenditures for home health care provided by informal care providers were assigned “-1” (INAPPLICABLE) because those types of events were skipped out (never asked) of the questions regarding expenditures.) “Household” edits were applied to sources and amounts of payment for all household-reported events for paid independent providers and unmatched agency providers. “MPC” edits were applied to provider-reported sources and amounts of payment for records matched to household-reported events for all agency home health providers. Both sets of edits were used to correct obvious errors in the reporting of expenditures. Imputations for independent paid providers and for agencies were conducted separately. Within each event type file, separate imputations were performed for flat fee and simple events.

Logical edits were used to sort each event into a specific category for the imputations. Events with complete expenditures were flagged as potential donors for the hot-deck imputations while events with missing expenditure data were assigned to various recipient categories. Each event with missing expenditure data was assigned to a recipient category based on the extent of its missing charge and expenditure data. For example, an event with a known total charge but no expenditure information was assigned to one category, while an event with a known total charge and partial expenditure information was assigned to a different category. Similarly, events without a known total charge and no or partial expenditure information were assigned to various recipient categories.

The logical edits produced eight recipient categories for HHP and eight recipient categories for HHA for events with missing data. Expenditures were imputed through separate hot-deck imputations for each of the eight recipient categories. The donor pool in these imputations includes events with complete expenditures from the HC for HHP or the MPC for HHA.

The donor pool included “free events” because, in some instances, providers are not paid for their services. These events represent charity care, bad debt, provider failure to bill, and third party payer restrictions on reimbursement in certain circumstances. (This does not include MPCELIG=3 (informal) events. As stated previously, home health care provided by friends, family, or volunteers (informal, MPCELIG=3) was assumed not to have expenditures associated with it and was not included in any imputation process.)

#### **2.6.4.3 Imputation Flag Variable (IMPFLAG)**

IMPFLAG is a six-category variable that indicates if the event contains complete Household Component (HC) or Medical Provider Component (MPC) data, was fully or partially imputed, or was imputed in the capitated imputation process. The following list identifies how the imputation flag is coded; the categories are mutually exclusive.

IMPFLAG=0 not eligible for imputation (includes zeroed out and flat fee leaf events)

IMPFLAG=1 complete HC data

IMPFLAG=2 complete MPC data

IMPFLAG=3 fully imputed

IMPFLAG=4 partially imputed

IMPFLAG=5 complete MPC data through capitation imputation (not applicable to HH)

#### **2.6.4.4 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 beginning in an earlier year, or (4) follow-up events were provided without a separate charge (e.g., after a surgical procedure). If all of the medical events for a person fell into one of these categories, then the total annual expenditures for that person would be zero. All expenditures for home health care provided by informal care providers (family, friends,

or volunteers, MPCELIG=3) were assigned -1 "INAPPLICABLE" because those types of events were skipped out (never asked) of the questions regarding expenditures.

#### **2.6.4.5 Sources of Payment**

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

1. Out-of-pocket by user or family,
2. Medicare,
3. Medicaid,
4. Private Insurance,
5. Veterans Administration, excluding TRICARE,
6. TRICARE,
7. Other Federal sources - includes Indian Health Service, Military Treatment Facilities, and other care by the Federal government,
8. Other State and Local sources - includes community and neighborhood clinics, State and local health departments, and State programs other than Medicaid,
9. Worker's Compensation, and
10. Other Unclassified sources - includes sources such as automobile, homeowner's, and liability insurance, and other miscellaneous or unknown sources.

Two additional source of payment variables were created to classify payments for events with apparent inconsistencies between insurance coverage and sources of payment based on data collected in the survey. These variables include:

11. Other Private - any type of private insurance payments reported for persons not reported to have any private health insurance coverage during the year as defined in MEPS, and
12. Other Public - Medicare/Medicaid payments reported for persons who were not reported to be enrolled in the Medicare/Medicaid program at any time during the year.

Though relatively small in magnitude, data users/analysts should exercise caution when interpreting the expenditures associated with these two additional sources of payment. While these payments stem from apparent inconsistent responses to health insurance and source of payment questions in the survey, some of these inconsistencies may 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 sampled 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 from persons who were not enrolled in Medicaid, but were presumed eligible by a provider who ultimately received payments from the public payer.

#### **2.6.4.6 Home Health Expenditure Variables (HHSF02X - HHXP02X)**

Home health agency, hospital, and nursing home events are sampled at a rate of 100% for the MPC. Households were not asked any expenditure-related questions in regards to these types of events; therefore, there are no household reported expenditure data for these events. Conversely, paid independent providers are not included in the MPC. Household reported responses are the only data available for these types of events. All expenditure data for paid independent providers are fully imputed from household reported expenditures. There are no expenditure data for informal care providers. Informal care (MPCELIG=3, unpaid care provided by family, friends, or volunteers) was assigned -1, “INAPPLICABLE”, in all expenditure categories.

The constructed variable MPCELIG is provided on this file. MPCELIG indicates whether the home health provider event was eligible for MPC data collection, and MPCELIG determines the imputation process applied to that event.

All of these expenditures have gone through an editing and imputation process and have been rounded to the nearest penny. HHSF02X - HHOT02X are the 12 sources of payment. HHTC02X is the total charge, and HHXP02X is the sum of the 12 sources of payment for the home health expenditures. The 12 sources of payment are: self/family (HHSF02X), Medicare (HHMR02X), Medicaid (HHMD02X), private insurance (HHPV02X), Veterans Administration (HHVA02X), TRICARE (HHTR02X), other Federal sources (HHOF02X), State and Local (non-federal) government sources (HHSL02X), Worker’s Compensation (HHWC02X), other private insurance (HHOR02X), other public insurance (HHOU02X), and other insurance (HHOT02X). Analysts can determine if a home health event was paid by an agency or some other paid independent provider by subsetting the variable MPCELIG to the appropriate and desired value.

#### **2.6.4.7 Rounding**

Expenditure variables on the 2002 home health event file have been rounded to the nearest penny. Person-level expenditure information released on the 2002 Person-Level Expenditure File was rounded to the nearest dollar. It should be noted that using the 2002 MEPS event files to create person-level totals will yield slightly different totals than those on the person-level expenditure file. These differences are due to rounding only. Moreover, in some instances, the number of persons having expenditures on the event files for a particular source of payment may differ from the number of persons with expenditures on the person-level expenditure file for that source of payment. This

difference is also an artifact of rounding only. Please see the MEPS 2002 Appendix File, HC-067I, for details on such rounding differences.

### **3.0 Sample Weight (PERWT02F)**

#### **3.1 Overview**

There is a single full year person-level weight (PERWT02F) assigned to each record for each key, in-scope person who responded to MEPS for the full period of time that he or she was in-scope during 2002. A key person either was a member of an NHIS household at the time of the NHIS interview, or became a member of 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 persons returning from military service, an institution, or living outside the United States). A person is in-scope whenever he or she is a member of the civilian noninstitutionalized portion of the U.S. population.

#### **3.2 Details on Person Weight Construction**

The person-level weight PERWT02F was developed in several stages. Person-level weights for Panels 6 and 7 were created separately. The weighting process for each panel included an adjustment for nonresponse over time and calibration to independent population figures. The calibration was initially accomplished separately for each panel by raking the corresponding sample weights to Current Population Survey (CPS) population estimates based on five variables. The five variables used in the establishment of the initial person-level control figures were: census region (Northeast, Midwest, South, West); MSA status (MSA, non-MSA); race/ethnicity (Hispanic, black but non-Hispanic, Asian but non-Hispanic and other); sex; and age. A 2002 composite weight was then formed by multiplying each weight from Panel 6 by the factor .55 and each weight from Panel 7 by the factor .45. The choice of factors reflected the relative sample sizes of the two panels, helping to limit the variance of estimates obtained from pooling the two samples. The composite weight was again raked to the same set of CPS-based control totals. When poverty status information derived from income variables became available, a final raking was undertaken on the previously established weight variable. Control totals were established using poverty status (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 the original five variables used in the previous calibrations.

### **3.2.1 MEPS Panel 6 Weight**

The person-level weight for MEPS Panel 6 was developed using the 2001 full year weight for an individual as a “base” weight for survey participants present in 2001. For key, in-scope respondents who joined an RU some time in 2002 after being out-of-scope in 2001, the 2001 family weight associated with the family the person joined served as a “base” weight. The weighting process included an adjustment for nonresponse over Rounds 4 and 5 as well as raking to population control figures for December 2002. These control figures were derived by scaling back the population totals obtained from the March 2002 CPS to reflect the December 2002 CPS estimated population distribution across age and sex categories as of December 2002. Variables used in the establishment of person-level control figures included: census region (Northeast, Midwest, South, West); MSA status (MSA, non-MSA); race/ethnicity (Hispanic, black but non-Hispanic, Asian but non-Hispanic, and other); sex; and age. Overall, the weighted population estimate for the civilian noninstitutionalized population on December 31, 2002 is 284,568,843. Key, responding persons not in-scope on December 31, 2002 but in-scope earlier in the year retained, as their final Panel 6 weight, the weight after the nonresponse adjustment.

### **3.2.2 MEPS Panel 7 Weight**

The person-level weight for MEPS Panel 7 was developed using the MEPS round 1 person-level weight as a “base” weight. For key, in-scope respondents who joined an RU after Round 1, the Round 1 family weight served as a “base” weight. The weighting process included an adjustment for nonresponse over Round 2 and the 2002 portion of Round 3 as well as raking to the same population control figures for December 2002 used for the MEPS Panel 6 weights. The same five variables employed for Panel 6 raking (census region, MSA status, race/ethnicity, sex, and age) were used for panel 7 raking. Similarly, for panel 7, key, responding persons not in-scope on December 31, 2002 but in-scope earlier in the year retained, as their final Panel 7 weight, the weight after the nonresponse adjustment.

Note that the MEPS Round 1 weights (for both panels with one exception as noted below) incorporated the following components: the original household probability of selection for the NHIS; ratio-adjustment to NHIS-based national population estimates at the household (occupied dwelling unit) level; adjustment for nonresponse at the dwelling unit level for Round 1; and raking to figures at the family and person level obtained from the March 2002 CPS data base.

### **3.2.3 The Final Weight for 2002**

Variables used in the establishment of person-level control figures included: poverty status (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); census



region (Northeast, Midwest, South, West); MSA status (MSA, non-MSA); race/ethnicity (Hispanic, black but non-Hispanic, Asian but non-Hispanic, and other); sex; and age. Overall, the weighted population estimate for the civilian noninstitutionalized population for December 31, 2002 is 284,568,843 (PERWT02F>0 and INSC1231=1). The weights of some persons out-of-scope on December 31, 2002 were also calibrated, this time using poststratification. Specifically, the weights of persons out-of-scope on December 31, 2002 who were in-scope some time during the year and also entered a nursing home during the year were poststratified to a corresponding control total obtained from the 1996 MEPS Nursing Home Component. The weights of persons who died while in-scope during 2002 were poststratified to corresponding estimates derived using data obtained from the Medicare Current Beneficiary Survey (MCBS) and Vital Statistics information provided by the National Center for Health Statistics (NCHS). Separate 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 is 288,181,763.

#### **3.2.4 Coverage**

The target population for MEPS in this file is the 2002 U.S. civilian noninstitutionalized population. However, the MEPS sampled households are a subsample of the NHIS households interviewed in 2000 (Panel 6) and 2001 (Panel 7). New households created after the NHIS interviews for the respective Panels and consisting exclusively of persons who entered the target population after 2000 (Panel 6) or after 2001 (Panel 7) 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. The set of uncovered persons constitutes only a small segment of the MEPS target population.

#### **4.0 Strategies for Estimation**

This file is constructed for efficient estimation of utilization, expenditures, and sources of payment for home health provider visits and to allow for estimates of number of persons with home health provider visits in 2002.

#### **4.1 Variables with Missing Values**

It is essential that the analyst examine all variables for the presence of negative values used to represent missing values. For continuous or discrete variables, where means or totals may be taken, it may be necessary to set minus values to values appropriate to the analytic needs. That is, the analyst should either impute a value or set the value to one that will be interpreted as missing by the computing language used. For categorical and dichotomous variables, the analyst may want to consider whether to recode or impute a

value for cases with negative values or whether to exclude or include such cases in the numerator and/or denominator when calculating proportions.

Methodologies used for the editing/imputation of expenditure variables (e.g., sources of payment, flat fee, and zero expenditures) are described in Section 2.6.4.2.

## 4.2 Basic Estimates of Utilization, Expenditures, and Sources of Payment

While the examples described below illustrate the use of event-level data in constructing person-level total expenditures, these estimates can also be derived from the person-level expenditure file unless the characteristic of interest is event specific.

In order to produce national estimates related to home health provider visits utilization, expenditures, and sources of payment, the value in each record contributing to the estimates must be multiplied by the weight (PERWT02F) contained on that record.

### Example 1

For example, the total number of home health paid independent provider visits, for the civilian noninstitutionalized population of the U.S. in 2002, is estimated as the sum of the weight (PERWT02F) across all home health paid independent provider records. That is,

$$\sum W_j = 5,036,044 \quad \text{across all records with MPCELIG} = 2 \quad (1)$$

### Example 2

Subsetting to records based on characteristics of interest expands the scope of potential estimates. For example, the estimate for the mean out-of-pocket payment per paid independent home health provider event (where the visit has a total expense greater than 0) should be calculated as the weighted mean of the paid independent home health provider's bill paid by self/family. That is,

$$(\sum W_j X_j) / (\sum W_j) = \$283.31 \quad (2)$$

where

$$\sum W_j = 4,571,718 \text{ and } X_j = \text{HHSF02X}_j$$

for all home health visits by paid independent provider (MPCELIG=2) with  $\text{HHXP02X}_j > 0$

This gives \$283.31 as the estimated mean amount of out-of-pocket payment of expenditures associated with home health events by paid independent providers and 4,571,718 as an estimate of the total number of home health events by paid independent

providers with expenditure. Both of these estimates are for the civilian noninstitutionalized population of the U.S. in 2002.

### Example 3

Another example would be to estimate the average proportion of total expenditures (where event expense is greater than 0) paid by private insurance for home health events by paid independent providers. This should be calculated as the weighted mean of the proportion of total expenditures paid by private insurance at the home health event-level. That is

$$(\sum W_j Y_j)/(\sum W_j) = 0.0245 \quad (3)$$

where

$$\sum W_j = 4,571,718 \text{ and } Y_j = \text{HHPV02X}_j / \text{HHXP02X}_j$$

for all home health visits by paid independent provider (MPCELIG=2) with  $\text{HHXP02X}_j > 0$

This gives 0.0245 as the estimated mean proportion of total expenditures paid by private insurance for home health events by paid independent providers with expenditures for the civilian noninstitutionalized population of the U.S. in 2002.

### 4.3 Estimates of the Number of Persons with Home Health Events Due to Hospitalization

When calculating an estimate of the total number of persons with home health events by paid independent providers, users can use a person-level file or this event file. However, this event file must be used when the measure of interest is defined at the event-level. For example, to estimate the number of home health events where services were provided due to hospitalization, this event file must be used. This would be estimated as,

$$\sum W_i X_i \quad \text{across all unique persons } i \text{ on this file} \quad (4)$$

where

$W_i$  is the sampling weight (PERWT02F) for person  $i$

and

$$X_i = 1 \quad \text{if HOSPITAL}_i = 1 \text{ for any home health visits by paid independent provider of person } i$$

$$= 0 \quad \text{otherwise.}$$

#### 4.4 Person-Based Ratio Estimates

##### 4.4.1 Person-Based Ratio Estimates Relative to Persons with Home Health Events by Paid Independent Providers

This file may be used to derive person-based ratio estimates. However, when calculating ratio estimates where the denominator is persons, care should be taken to properly define and estimate the unit of analysis up to person-level. For example, the mean expense for persons with home health events by paid independent providers (MPCELIG =2) is estimated as,

$$\frac{(\sum W_i Z_i)}{(\sum W_i)} \quad \text{across all unique persons } i \text{ on this file} \quad (5)$$

where

$W_i$  is the sampling weight (PERWT02F) for person  $i$

and

$Z_i = \sum \text{HHXP02X}_i$  across all home health visits by paid independent providers for person  $i$ .

##### 4.4.2 Person-Based Ratio Estimates Relative to the Entire Population

If the ratio relates to the entire population, this file cannot be used to calculate the denominator, as only those persons with at least one home health provider event are represented on this data file. In this case, the Full Year Consolidated, which has data for all sampled persons, must be used to estimate the total number of persons (i.e., those with events and those without events).

For example, to estimate the proportion of the civilian noninstitutionalized population of the U.S. with at least one home health event by a paid independent provider, the numerator would be derived from data on this event file, and the denominator would be derived from data on the person-level file. That is,

$$\frac{(\sum W_i Z_i)}{(\sum W_i)} \quad \text{across all unique persons } i \text{ on the MEPS HC FY person-level file} \quad (6)$$

where

$W_i$  is the sampling weight (PERWT02F) for person  $i$

and

$Z_i = 1$  if MPCELIG <sub>$i$</sub>  = 2 for any home health visits by paid independent provider of person  $i$   
 $= 0$  otherwise.

#### **4.5 Sampling Weights for Merging Previous Releases of MEPS Household Data with this Event File**

There have been several previous releases of MEPS Household Survey public use data. Unless a variable name common to several files is provided, the sampling weights contained on these data files are file-specific. The file-specific weights reflect minor adjustments to eligibility and response indicators due to birth, death, or institutionalization among respondents.

For estimates from a MEPS data file that do not require merging with variables from other MEPS data files, the sampling weight(s) provided on that data file are the appropriate weight(s). When merging one MEPS Household data file to another, the major analytical variable (i.e., the dependent variable) determines the correct sampling weight to use.

#### **4.6 Variance Estimation (VARPSU, VARSTR)**

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. Various approaches can be used to develop such estimates of variance including use of the Taylor Series or various replication methodologies. Replicate weights have not been developed for the MEPS 2002 data. Variables needed to implement a Taylor Series estimation approach are provided in the file and are described in the paragraph below.

Using a Taylor Series approach, variance estimation strata and the variance estimation PSUs within these strata must be specified. The corresponding variables on the MEPS full year utilization database are VARSTR and VARPSU, respectively. 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. However, beginning with the 2002 Point-in-Time PUF, the variance strata and PSUs have been developed to be compatible with all future PUFs. Thus, data from future years can be pooled and 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. Specifying a “with replacement” design in a computer software package such as SUDAAN (Shah, 1996) should provide 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 actual number available. For MEPS sample estimates for characteristics generally distributed throughout the country (and thus the sample PSUs), there are over 100 degrees of freedom associated with the corresponding estimates of variance. The following illustrates these concepts using two examples from section 4.2.

## **Examples 2 and 3 from Section 4.2**

Using a Taylor Series approach, specifying VARSTR and VARPSU as the variance estimation strata and PSUs (within these strata) respectively and specifying a “with replacement” design in a computer software package (i.e., SUDAAN will yield standard error estimates of \$53.76 and 0.0111 for the estimated mean of out-of-pocket payment and the estimated mean proportion of total expenditures paid by private insurance respectively.

### **5.0 Merging/Linking MEPS Data Files**

Data from this file can be used alone or in conjunction with other files. This section provides instructions, or the details on where to find the instructions, for linking the 2002 home health provider events with other 2002 MEPS public use files, including the 2002 conditions file and a 2002 person-level file.

#### **5.1 Linking a 2002 Person-Level File to the 2002 Home Health Event File**

Merging characteristics of interest from other 2002 MEPS files (e.g., the 2002 Full Year Population Characteristics File or the 2002 Prescribed Medicines File) expands the scope of potential estimates. For example, to estimate the total number of home health provider events of persons with specific characteristics (e.g., age, race, and sex), population characteristics from a person-level file need to be merged onto the home health provider file. This procedure is illustrated below. The MEPS 2002 Appendix File, HC-067I, provides additional details on how to merge 2002 MEPS data files.

1. Create data set PERSX by sorting the 2002 Full Year Population Characteristics File by the person identifier, DUPERSID. Keep only variables to be merged on to the home health provider event file and DUPERSID.
2. Create data set HVIS by sorting the home health provider event file by person identifier, DUPERSID.
3. Create final data set NEWHVIS by merging these two files by DUPERSID, keeping only records on the home health provider event file.

The following is an example of SAS code, which completes these steps:

```
PROC SORT DATA=HCXXX(KEEP=DUPERSID AGE31X AGE42X
AGE53X SEX RACEX EDUCYR) OUT=PERSX;
  BY DUPERSID;
RUN;

PROC SORT DATA=HVIS;
  BY DUPERSID;
RUN;

DATA NEWHVIS;
  MERGE HVIS (IN=A) PERSX(IN=B);
  BY DUPERSID;
  IF A;
RUN;
```

## **5.2 Linking the 2002 Home Health Event File to the 2002 Medical Conditions File**

Due to survey design issues, there are limitations/caveats that data users/analysts must keep in mind when linking the different files. These limitations/caveats are listed below. For detailed linking examples, including SAS code, data users/analysts should refer to the MEPS 2002 Appendix File, HC-067I.

### **5.2.1 Limitations/Caveats of CLNK (the 2002 Medical Conditions Link File)**

The CLNK provides a link from 2002 MEPS event files to the 2002 Medical Conditions File. When using the CLNK, data users/analysts should keep in mind that (1) conditions are self-reported and (2) there may be multiple conditions associated with a home health provider event. Data users/analysts should also note that not all home health provider events link to the condition file.

## References

- Cohen, S.B. (1999). Sample Design of the 1996 Medical Expenditure Panel Survey Medical Provider Component. Journal of Economic and Social Measurement. Vol 24, 25-53.
- Cohen, S.B. (1997). Sample Design of the 1996 Medical Expenditure Panel Survey Household Component. Rockville (MD): Agency for Health Care Policy and Research; 1997. *MEPS Methodology Report, No. 2*. AHCPR Pub. No. 97-0027.
- Cohen, J.W. (1997). Design and Methods of the Medical Expenditure Panel Survey Household Component. Rockville (MD): Agency for Health Care Policy and Research; 1997. *MEPS Methodology Report, No. 1*. AHCPR Pub. No. 97-0026.
- Cohen, S.B. (1996). The Redesign of the Medical Expenditure Panel Survey: A Component of the DHHS Survey Integration Plan. *Proceedings of the COPAFS Seminar on Statistical Methodology in the Public Service*.
- Cox, B.G. and Cohen, S.B. (1985). Chapter 8: Imputation Procedures to Compensate for Missing Responses to Data Items. In *Methodological Issues for Health Care Surveys*. Marcel Dekker, New York.
- Monheit, A.C., Wilson, R., and Arnett, III, R.H. (Editors) (1999). *Informing American Health Care Policy*. Jossey-Bass Inc, San Francisco.
- Shah, B.V., Barnwell, B.G., Bieler, G.S., Boyle, K.E., Folsom, R.E., Lavange, L., Wheelless, S.C., and Williams, R. (1996). *Technical Manual: Statistical Methods and Algorithms Used in SUDAAN Release 7.0*, Research Triangle Park, NC: Research Triangle Institute.



#### **D. VARIABLE-SOURCE CROSSWALK**

## VARIABLE-SOURCE CROSSWALK

### FOR MEPS HC-067H: 2002 HOME HEALTH VISITS

#### Survey Administration Variables

Variable	Description	Source
DUID	Dwelling unit ID	Assigned in sampling
PID	Person number	Assigned in sampling
DUPERSID	Person ID (DUID + PID)	Assigned in sampling
EVNTIDX	Event ID	Assigned in sampling
EVENTRN	Event round number	CAPI derived

#### Home Health Events Variables

Variable	Description	Source
HHDATEYR	Event date – year	CAPI derived
HHDATEMM	Event date – month	CAPI derived
MPCELIG	MPC eligibility flag	Constructed
SELFAGEN	Does provider work for agency or self	EV06A
HHTYPE	Home health event type	EV06
CNA	Type of health care worker – certified nurse assistant	HH01
COMPANN	Type of health care worker – companion	HH01
DIETICN	Type of health care worker – dietitian/nutritionist	HH01
HHAIDE	Type of health care worker – home care aide	HH01
HOSPICE	Type of health care worker – hospice worker	HH01
HMEMAKER	Type of health care worker - homemaker	HH01
IVTHP	Type of health care worker – IV therapist	HH01
MEDLDOC	Type of health care worker – medical doctor	HH01

<b>Variable</b>	<b>Description</b>	<b>Source</b>
NURPRACT	Type of health care worker – nurse/practitioner	HH01
NURAIDE	Type of health care worker – nurse’s aide	HH01
OCCUPTHP	Type of health care worker – occupational therapist	HH01
PERSONAL	Type of health care worker – personal care attendant	HH01
PHYSLTHP	Type of health care worker – physical therapy	HH01
RESPTHP	Type of health care worker – respiratory therapy	HH01
SOCIALW	Type of health care worker – social worker	HH01
SPEECTHP	Type of health care worker – speech therapy	HH01
OTHRHCW	Type of health care worker – other	HH01
NONSKILL	Type of health care worker – non-skilled	HH02
SKILLED	Type of health care worker – skilled	HH02
SKILLWOS	Specify type of skilled worker	HH02OV1
OTHCW	Type of health care worker – some other	HH02
OTHCWOS	Specify other type health care worker	HH02OV2
HOSPITAL	Any home health care service due to hospitalization	HH03
VSTRELCN	Any Home Health Care Service Related to Health Condition	HH04
TREATMT	Person received medical treatment	HH06
MEDEQUIP	Person was taught use of medical equipment	HH07
DAILYACT	Person was helped with daily activities	HH08
COMPANY	Person received companionship services	HH09
OTHSVCE	Person received other home care services	HH10
OTHSVCOS	Specify other home care services received	HH10OV
FREQCY	Provider helped every week/some weeks	HH11
DAYSPWK	# days per week provider came (agency only)	HH12
DAYSPMO	# days per month provider came (agency only)	HH13
HOWOFTEN	Provider came once per day/more than once per day	HH14
TMSPDAY	Times/day provider came to home to help	HH15
HRSLONG	Hours each visit lasted	HH16

<b>Variable</b>	<b>Description</b>	<b>Source</b>
MINLONG	Minutes each visit lasted	HH16
SAMESVCE	Any other months person received same services	HH17
HHDAYS	Days per month in home health, 2002	Constructed

### **Imputed Expenditure Variables**

<b>Variable</b>	<b>Description</b>	<b>Source</b>
HHSF02X	Amount paid, self/family (Imputed)	CP Section (Edited)
HHMR02X	Amount paid, Medicare (Imputed)	CP Section (Edited)
HHMD02X	Amount paid, Medicaid (Imputed)	CP Section (Edited)
HHPV02X	Amount paid, private insurance (Imputed)	CP Section (Edited)
HHVA02X	Amount paid, Veterans Administration (Imputed)	CP Section (Edited)
HHTR02X	Amount paid, TRICARE (Imputed)	CP Section (Edited)
HHOF02X	Amount paid, other federal (Imputed)	CP Section (Edited)
HHSL02X	Amount paid, state & local government (Imputed)	CP Section (Edited)
HHWC02X	Amount paid, workers' compensation (Imputed)	CP Section (Edited)
HHOR02X	Amount paid, other private insurance (Imputed)	Constructed
HHOU02X	Amount paid, other public insurance (Imputed)	Constructed
HHOT02X	Amount paid, other insurance (Imputed)	CP Section (Edited)
HHXP02X	Sum of HHSF02X – HHOT02X (Imputed)	Constructed
HHTC02X	Household reported total charge (Imputed)	CP Section (Edited)
IMPFLAG	Imputation status	Constructed

### **Weights**

<b>Variable</b>	<b>Description</b>	<b>Source</b>
PERWT02F	Expenditure file person weight, 2002	Constructed
VARSTR	Variance estimation stratum, 2002	Constructed
VARPSU	Variance estimation PSU, 2002	Constructed