MEPS HC-026H:
1998 Home Health File

October 2001

Agency for Healthcare Research and Quality
Center for Cost and Financing Studies

Table of Contents

A. Data Use Agreement

Individual identifiers have been removed from the microdata contained in the files on this CD-ROM. 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.

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.

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 18 U.S.C. 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

This documentation describes one in a series of public use files from the Medical Expenditure Panel Survey (MEPS). The survey provides a new and extensive data set on the use of health services and health care in the United States.

MEPS is conducted to provide nationally representative estimates of health care use, expenditures, sources of payment, and insurance coverage for the U.S. civilian noninstitutionalized population. MEPS also includes a nationally representative survey of nursing homes and their residents. MEPS is cosponsored by the Agency for Healthcare Research and Quality (AHRQ) (formerly the Agency for Health Care Policy and Research (AHCPR)) and the National Center for Health Statistics (NCHS).

MEPS comprises three component surveys: the Household Component (HC), the Medical Provider Component (MPC), and the Insurance Component (IC). The HC is the core survey, and it forms the basis for the MPC sample and part of the IC sample. The separate NHC sample supplements the other MEPS components. 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. The National Medical Expenditure Survey (NMES-2) was conducted in 1987. Beginning in 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 system.

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 accommodate these goals, new MEPS design features include linkage with the National Health Interview Survey (NHIS), from which the sampling frame for the MEPS HC is drawn, and continuous 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 2½ -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, conducted by NCHS. 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 validates 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 hospitals, hospital physicians, home health agencies, and pharmacies reported in the HC. Also included in the MPC are all office-based physicians who:

- were identified by the household respondent as providing care for HC respondents receiving Medicaid.

- were selected through a 75-percent sample of HC households receiving care through an HMO (health maintenance organization) or managed care plan.

- were selected through a 25-percent sample of the remaining HC households.

Data are collected on medical and financial characteristics of medical and pharmacy events reported by HC respondents, including:

- Diagnoses coded according to ICD-9-CM (9th Revision, International Classification of Diseases) and DSM-IV (Fourth Edition, *Diagnostic and Statistical Manual of Mental Disorders*).

- Physician procedure codes classified by CPT-4 (Common Procedure Terminology, Version 4).

- Inpatient stay codes classified by DRGs (diagnosis-related groups).

- Prescriptions coded by national drug code (NDC), medication name, strength, and quantity dispensed.

- Charges, payments, and the reasons for any difference between charges and payments.

The MPC is conducted through telephone interviews and mailed survey materials. In some instances, providers sent medical and billing records which were abstracted into the survey instruments.

3.0　　Insurance Component

The MEPS IC collects data on health insurance plans obtained through employers, unions, and other sources of private health insurance. 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 four 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.

- An Internal Revenue Service list of the self-employed.

To provide an integrated picture of health insurance, data collected from the first sampling frame (employers and insurance providers) are linked back to data provided by the MEPS HC respondents. Data from the other three sampling frames are collected to provide annual national and State estimates of the supply of private health insurance available to American workers and to evaluate policy issues pertaining to health insurance.

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 and microdata files. Summary reports are released as printed documents and electronic files. Microdata files are released on CD-ROM and/or as electronic files.

Printed documents and CD-ROMs 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 or CD-ROM you are requesting. Selected electronic files are available from the Internet on the MEPS web site: <http://www.meps.ahrq.gov/>.

Additional information on MEPS is available from the MEPS project manager or the MEPS public use data manager at the Center for Cost and Financing Studies, Agency for Healthcare Research and Quality.

C. Technical and Programming Information

1.0 General Information

This documentation describes one in a series of public use event files from the 1998 Medical Expenditure Panel Survey Household (HC) and Medical Provider Components (MPC). Released as an ASCII data file and SAS transport file, this public use file provides detailed information on home health events for a nationally representative sample of the civilian noninstitutionalized population of the United States and can be used to make estimates of home health utilization and expenditures for calendar year 1998. As illustrated below, this file consists of MEPS survey data obtained in the 1998 portion of round 3, and rounds 4 and 5 for Panel 2, as well as rounds 1, 2, and the 1998 portion of round 3 for Panel 3 of the MEPS HC (i.e., the rounds for the MEPS panels covering calendar year 1998).
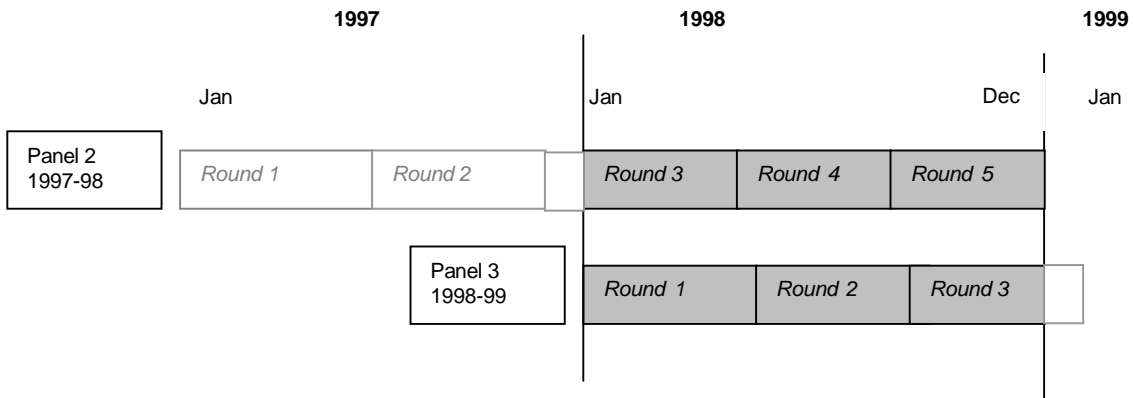
Counts (utilization) of home health events are based entirely on household reports. Agency home health providers were sampled into the MEPS MPC (see Section B. 2.0). Only those providers for whom the respondent signed a permission form were included in MPC. Information from MPC was used to supplement expenditure and payment data reported by the household.

Data from this event file can be merged with other 1998 MEPS HC data files for purposes of appending person characteristics, such as demographic or health insurance coverage to each home health record.

This file can be also used to construct summary variables of expenditures, sources of payment, and related aspects of home health events for calendar year 1998. Aggregate annual person-level information on the use of home health providers and other health services use is provided on the 1998 Population Characteristics file, where each record represents a MEPS sampled person. However, the 1998 Population Characteristics File contained preliminary utilization estimates that included duplicate records. These duplicate records have been eliminated from this event file and will be deleted from MEPS: HC-028 1998 Person-level Expenditures and Utilization File. Therefore, utilization counts on the Person-level Expenditures and Utilization File, and the 1998 Population Characteristics File will not match.

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

> Data File Information
> Sample Weights
> Merging MEPS Data Files
> References
> Variable to 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, 1998. 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

This public use data set consists of two event-level data files. File 1 contains characteristics associated with the home health event and imputed expenditure data. File 2 contains unimputed expenditure data from both the Household and Medical Provider Components for all home health events on File 1 except for those records representing care from informal care providers, for which no expenditure information was obtained.

Each record represents a household-reported home health event. A home health event is a MONTH of similar service 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 4 events from a nurse, 10 events from a homemaker and 4 events from a physical therapist all from the same provider every month for 3 months, then there will be 3 event records on the file, one for each month (NOT 54 records). 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 providers. Persons with more than one event are represented on this file more than once. Likewise, persons who do not have a home health event are not represented on the file.

File 1 of this public use data set contains 3,904 home health records. Of the 3,904 records, 3,839 are associated with persons having a positive person-level weight (WTDPER98). File 1 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. File 2 does not include those records in which the care received was from an informal care provider. Each record represents one household-reported home health event that occurred during calendar year 1998. 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 1998 portion of round 3, and rounds 4 and 5 for Panel 2, as well as rounds 1, 2, and the 1998 portion of round 3 for Panel 3 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 1998 eligibility (i.e., persons with a positive 1998 full-year person-level sampling weight (WTDPER98>0)), or

(b) Be classified as either an eligible non-key person or an eligible out-of-scope person who responded for his or her entire period of 1998 eligibility, and belonged to a family (i.e., all persons with the same value for FAMID) in which all eligible family members responded for their entire period of 1998 eligibility, and at least one family member has a positive 1998 full-year person weight (i.e., eligible non-key or eligible out-of-scope persons who are members of a family all of whose members have a positive 1998 full-year MEPS family-level weight (WTFAM98>0)).

Please refer to Attachment 1 for definitions of key, non-key, inscope and eligible. Persons with no home health events for 1998 are not included on this file (but are represented on the 1998 MEPS person-level files). A codebook for each of the data files is provided in files H26HF1CB.PDF and H26HF2.CB.PDF.

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

For home health agencies, hospitals, and nursing homes, 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, hospitals, and nursing homes. Paid independent providers generally include companions, nursing assistants, physicians, etc. For each record on File 1, one or more types of workers can be reported. The respondent is asked to mention all of the types of home health workers whom 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 event 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 on File 1 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; if care was received due to hospitalization; whether or not a person was taught how to use medical equipment;

flat fee information; imputed sources of payment, total payment and total charge for the home health event expenditure; and a full-year person-level weight.

File 2 of this public use data set contains 3,320 home health records. File 2 has less records than File 1 because home health records where friends, family and volunteers provided the home health care to a person are considered unpaid and are not included on File 2 (no expenditure information is available for them). File 2 of this public use data set is intended for analysts who want to perform their own imputations to handle missing data. This file contains one set of unimputed expenditure information from the MPC (if home health provider was sampled in the MPC), as well as one set of pre-imputed expenditure information from the HC. Both sets of expenditure data have been subject to minimal logical editing that accounted for outliers, copayments or charges reported as total payments, and reimbursed amounts that were reported as out of pocket payments. In addition, edits were implemented to correct for mis-classifications between Medicare and Medicaid and between Medicare HMO's and private HMO's as payment sources. However, missing data were not imputed.

Data from these files can be merged with previously released 1998 MEPS HC person-level data using the unique person identifier, DUPERSID, to append person-level characteristics such as demographic or health insurance coverage to each record. The 1998 home health event file can also be linked to the MEPS 1998 Medical Conditions File and the MEPS 1998 Prescribed Medicines File. Please see Section 5.0 and the 1998 Appendix File for details on how to link MEPS data files.

Panel 2 cases (Panel98=2 on the 1998 person-level file) can be linked back to the 1997 MEPS HC Public Use Data Files. However, the user should be aware, at this time, no weight is being provided to facilitate 2-year analysis of Panel 2 data.

2.1   Codebook Structure

For each variable on these files, both weighted and unweighted frequencies are provided in the codebooks (in files H26HF1CB.PDF and H26HF2CB.PDF). The codebook and data file sequence list variables in the following order:

File 1
        Unique person identifiers
        Unique home health event identifier
        Other survey administration variables
        Home health characteristic variables
        Imputed expenditure variables
        Weight and variance estimation variables

File 2
        Unique person identifiers
        Unique home health event identifier
        Pre-imputed expenditure variables
        Weight and variance estimation variables

2.2   Reserved Codes

The following reserved code values are used:

Value	Definition
-1 INAPPLICABLE	Question was not asked due to skip pattern.
-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, -1,-7, -8, and -9 have not been edited on this file. The values of -1 and -9 can be edited by analysts by following the skip patterns in the questionnaire. A copy of the Household Component questionnaire can be found on the World Wide Web at http://meps.ahrq.gov and clicking on the link in the Home Health box.

2.3   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:

Identifier	Description
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

2.4   Variable Naming

In general, variable names reflect the content of the variable, with an 8-character limitation. All imputed/edited variables end with an "X."

2.4.1   General

Variables contained on Files 1 and 2 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, entitled, "Variable - Source Crosswalk." Sources for each variable are indicated in one of four ways: (1) variables which are derived from CAPI or assigned in sampling are so indicated; (2) variables which come from one or more specific questions have those numbers and the questionnaire section

indicated in the "Source" column; (3) variables constructed from multiple questions using complex algorithms are labeled "Constructed" in the "Source" column; and (4) variables which have been imputed are so indicated.

2.4.2   Expenditure and Sources of Payment Variables

The pre-imputed and imputed versions of the expenditure and sources of payment variables are provided on the 2 separate files. Variables on Files 1 and 2 follow a standard naming convention and are 8 characters in length. Please note that pre-imputed means that a series of logical edits have been performed on the variable but missing data remains. The imputed versions incorporate the same edits but have also undergone an imputation process to account for missing data.

All imputed variables on File 1 end with an "X". The pre-imputed expenditure variables on File 2 end with an "H".

The total sum of payments, 12 sources of payment variables and total charge variables are named consistently 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 sources of payment variables, the third and fourth characters indicate:

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

The fifth and sixth characters indicate the year (98). The last character indicates whether it is edited/imputed (X) or is a pre-imputed variable (H).

For example, HHSF98X is the edited/imputed amount paid by self or family for a home health event expenditure incurred in 1998.

2.5 File 1 Contents

2.5.1 Survey Administration

2.5.1.1 Person Identifiers (DUID, PID, DUPERSID)

The dwelling unit ID (DUID) is a 5-digit random number assigned after the case was sampled for MEPS. The 3-digit person number (PID) uniquely identifies each person within the dwelling unit. The 8-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 Attachment 1.

2.5.1.2 Record Identifiers (EVNTIDX, FFEEIDX, EVENTRN)

EVNTIDX uniquely identifies each event (i.e., each record on the file).

FFEEIDX uniquely identifies a flat fee group, that is, all events that were part of a flat fee payment situation. For example, pregnancy is typically covered in a flat fee arrangement where the prenatal visits, the delivery, and the postpartum visits are all covered under one flat fee dollar amount. These three events (the prenatal visit, the delivery, and the postpartum visits) have the same value for FFEEIDX. Please note that FFEEIDX should be used to link up all 1998 MEPS event files (excluding prescribed medicines) in order to determine the full set of events that are part of a flat fee group. *Although four households reported home health events initially had valid flat fee identifiers (HHSFFEID), they were all disavowed by MPC data, and therefore, all values of FFEDIDX were set to –1 on File 1.*

EVENTRN indicates the round in which the home health event was first reported.

2.5.2 Characteristics of Home Health Events

File 1 contains 44 variables describing home health events reported by respondents in the Home Health section of the MEPS-HC questionnaire. The questionnaire contains specific probes for determining specific details about the home health event.

2.5.2.1 Date Home Health Event Started (HHBEGYR, HHBEGMM)

The start date variables (HHBEGYR and HHBEGMM) 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 HHBEGYR is that all events are reported as having started in 1998 even though a person could have started receiving that type of home health care from that provider year(s) before 1998. These variables should not be interpreted as "true" start dates.

2.5.2.2 Characteristics of Home Health Events (MPCELIG-OTHCWOS)

The HC questionnaire determines whether the home health provider event(s) for each month's services was an agency or whether the provider was an independent paid provider (SELFAGEN). Respondents were also asked if the provider was paid or whether services were provided by a friend, relative, or volunteer (HHTYPE). 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, the data were not edited to ensure consistency between the variables MPCELIG, SELFAGEN, and HHTYPE for all home health provider events. If necessary, analysts are free to edit these variables as they see fit. All respondents receiving care from an agency, hospital or nursing home were asked to identify the type of home health worker they saw (CNA-SPEECHTHP) -- 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. To make estimates of personal care attendants or home care aides based on the their identification by household respondents and by 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 and skilled 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-SPEECHTHP). In addition, the type of health care worker variables (CNA-SPEECHTHP) were not reconciled with MPCELIG, SELFAGEN and HHTYPE, so inconsistencies between these variables remain in the data.

2.5.2.3 Treatments, Therapies and Services (HOSPITAL-OTHSVCOS)

Regardless of the type of provider, all respondents were asked if the home health services they received were due to a hospitalization (HOSPITAL), whether it was 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, hospital, or nursing home were asked if they were taught how to use medical equipment (MEDEQUIP) and whether or not they received a medical treatment (TREATMT).

2.5.2.4 Frequency of Home Health Events (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 respondent 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 used in conjunction with HRSLONG and TMSPDAY, can be used to form a measure of intensity of care -- that is, how many hours of care was provided in one month.

2.5.3 Condition and Procedure Codes 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 1998 Medical Conditions File. Details on how to link to the MEPS 1998 Medical Conditions File are provided in the 1998 Appendix File.

2.5.4 Flat Fee Variables

User's Note: For home health events, use flat fee variables with caution. Flat fees are not common with respect to home health events (***there are no home health events on File 1 that are considered flat fee events***) and should not be a focus of an analysis.

2.5.4.1 Definition of Flat Fee Payments

A flat fee is the fixed dollar amount a person is charged for a package of health care services. An example is obstetrician's fee covering a normal delivery, as well as pre- and post-natal care. A flat fee group is the set of medical services (i.e., events) that are covered under the same flat fee payment situation. The flat fee groups represented on 1998 MEPS event files, include flat fee groups where at least one of the health care events, as reported by the HC respondent, occurred during 1998. By definition a flat fee group can span multiple years and/or event types (e.g., hospital stay, physician office visit), and a single person can have multiple flat fee groups.

2.5.4.2 Flat Fee Variable Descriptions

FFHHTYPE indicates whether the 1998 home health provider event is the "stem" or "leaf" of a flat fee group. A stem (records with FFHHTYPE = 1) is the initial medical service (event) which is followed by other medical events that are covered under the same flat fee payment. The leaf of the flat fee group (records with FFHHTYPE = 2) are those medical events that are tied back to the initial medical event (the stem) in the flat fee group.

Please note, there are no home health events on File 1 that are considered flat fee events; therefore, all events will have –1 for FFHHTYPE.

2.5.4.3 Counts of Flat Fee Events that Cross Years (FFBEF98 – FFTOT99)

As described above, a flat fee payment situation covers multiple events and the multiple events could span multiple years. For situations where a 1998 home health provider event is part of a group of

events, and some of the events occurred before 1998, counts of the known events are provided on the home health provider event file record. An indicator variable is provided if some of the events occurred after 1998. These variables are:

FFBEF98 -- total number of pre-1998 events in the same flat fee group as the 1998 home health provider event record. This count would not include 1998 home health provider event.

FFTOT99 -- indicates whether or not there any 1999 medical events in the same flat fee group as the 1998 home health provider event record.

Please note, there are no home health events on File 1 that are considered flat fee events; therefore, all events will have –1 for FFBEF98 and FFTOT99.

2.5.4.4   Caveats of Flat Fee Groups

The user should note that flat fee payment situations are not common with respect to home health provider events. *There are no home health events on File 1 that are considered flat fee events.*

In general, every flat fee group should have an initial event (stem) and at least one subsequent event (leaf). There are some situations where this is not true. For some of these flat fee groups, the initial event reported occurred in 1998 but the remaining events that were part of this flat fee group occurred in 1999. In this case, the 1998 flat fee group represented on this file would consist of one event (the stem). The 1999 events that are part of this flat fee group are not represented on this file. Similarly, the household respondent may have reported a flat fee group where the initial event began in 1997 but subsequent events occurred during 1998. In this case, the initial event would not be represented on the file. This 1998 flat fee group would then only consist of one or more leaf records and no stem. Another reason for which a flat fee group would not have a stem and a leaf record is that the stems or leaves could have been reported as different event types.

2.5.5   Expenditure Data

2.5.5.1   Definition of Expenditures

Expenditures on Files 1 and 2 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. For details on expenditure definitions, please reference the following, "Informing American Health Care Policy" (Monheit et al., 1999). AHRQ has developed factors to apply to the 1987 NMES expenditure data to facilitate longitudinal analysis. These factors can be accessed via the CCFS Data Center. For more information, see the Data Center section of the MEPS Web Site at http://www.meps.ahrq.gov.

2.5.5.2   Data Editing/Imputation Methodologies of Expenditure Variables

General Imputation Methodology

The general methodology used for editing and imputing expenditure data is described below. However, please note, home health events provided by an agency, hospital or nursing home were included in the MPC, and home health provided by paid independent providers were not included in the MPC. 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 were assumed to be free and were 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.

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.

A weighted sequential hot-deck procedure was used to impute for missing expenditures, as well as total charge. The 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.

Imputation Methodology for Home Health Events

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 were assumed to not have expenditures associated with them and were 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 of (never asked) the questions regarding expenditures. "Household" edits were applied to sources and amounts of payment for all events reported for paid independent providers by HC respondents. "MPC" edits were applied to provider-reported sources and amounts of payment for records matched to household-reported events for all agency, hospital, and nursing home 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, hospitals, and nursing homes were conducted separately. Separate imputations also were performed for 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 was assigned to a recipient category based on its pattern of missing 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 some expenditure information was assigned to a different category. Similarly, events without a known total charge were assigned to various recipient categories based on the amount of missing data.

The logical edits produced eight recipient categories 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 was restricted to events with complete expenditures from either the HC or the MPC.

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. If free events were excluded from the donor pool, total expenditures would be over-counted because the cost of free care would be implicitly included in paid events and explicitly included in events that should have been treated as free from provider. Analysts should note that home health care provided by friends, family, or volunteers (informal, MPCELIG=3) were assumed to not have expenditures associated with them and were not included in any imputation process.

Flat Fee Expenditures

The approach used to count expenditures for flat fees was to place the expenditure on the first event of the flat fee group. The remaining events have zero payments. Thus, if the first event in the flat fee group occurred prior to 1998, all of the events that occurred in 1998 will have zero payments. Conversely, if the first event in the flat fee group occurred at the end of 1998, the total expenditure for the entire flat fee group will be on that event, regardless of the number of events it covered after 1998.

Please note, there are no home health events on File 1 that are considered flat fee events.

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 of (never asked) the questions regarding expenditures.

Discount Adjustment Factor

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

Sources of Payment

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

1. Out of pocket by user or family
2. Medicare
3. Medicaid
4. Private Insurance
5. Veteran's Administration, excluding CHAMPVA
6. CHAMPUS or CHAMPVA
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
10. Other Unclassified sources - includes sources such as automobile, homeowner's, liability, and other miscellaneous or unknown sources.

Two additional sources 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 - Medicaid payments reported for persons who were not reported to be enrolled in the Medicaid program at any time during the year.

Though relatively small in magnitude, users 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 sources 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 persons who were not enrolled in Medicaid, but were presumed eligible by a provider who ultimately received payments from the program.

Users should also note that the Other Public and Other Private sources of payment categories only exist on File 1 for imputed expenditure data since they were created through the editing/imputation process. File 2 reflects 10 sources of payment as they were collected through the MEPS HC and MPC survey instruments.

Home Health Expenditure Variables (HHSF98X - HHXP98X and HHSF98H-HHUC98H)

There are 12 expenditure variables specific to paid independent home health events (MPCELIG=2) and 14 expenditure variables specific to agency home health events (MPCELIG=1). 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) were assigned -1 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. Users should be aware that MPCELIG was not reconciled with SELFAGEN and there may be inconsistencies between the two variables.

All of these expenditures have gone through an editing and imputation process and have been rounded to the second decimal place. There is a sum of payments variable (HHXP98X) which for each home health event sums all the expenditures from the various sources of payment. The 12 sources of payment expenditure variables for each home health event are the following: amount paid by self or family (HHSF98X), amount paid by Medicare (HHMR98X), amount paid by Medicaid (HHMD98X), amount paid by private insurance (HHPV98X), amount paid by Veteran's Administration (HHVA98X), amount paid by CHAMPUS/CHAMPVA (HHCH98X), amount paid other Federal sources (HHOF98X), amount paid by State and Local (non-federal) government sources (HHSL98X), amount paid by Worker's Compensation (HHWC98X), and amount paid by some other source of insurance (HHOT98X). As mentioned previously, there are two additional expenditure variables called HHOR98X and HHOU98X (other private and other public, respectively). These two expenditure variables were created to maintain consistency between what the household reported as their private and public insurance status for hospitalization and physician coverage. 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.

Rounding

Expenditure variables on the 1998 home health event file have been rounded to the nearest penny. Person-level expenditure information released on the 1998 person-level expenditure file were rounded to the nearest dollar. It should be noted that using the 1998 MEPS event files to create person-level totals will yield slightly different totals than those on the 1998 person-level expenditure file. These differences are due to rounding only. Moreover, in some instances, the number of persons having expenditures on the 1998 event files for a particular source of payment may differ from the number of persons with expenditures on the 1998 person-level expenditure file for that source of payment. This difference is also an artifact of rounding only. Please see the 1998 Appendix File for details on such rounding differences.

Identifying Imputed Expenditures

If the user desires to identify whether sources of payment and total charge have been imputed, simply compare the expenditure variable of interest from File 2 with the corresponding variable from File 1. An imputed value would be one having a missing value on File 2 while the value on File 1 would be zero or greater. In a small number of cases, an imputed value on File 1 will have a corresponding value of zero rather than missing on File 2.

The user should note that there are 10 sources of payment variables in the pre-imputed expenditure data on File 2, while the imputed expenditure data on File 1 contains 12 sources of payment variables. The additional two sources of payment (which are not reported as separate sources of payment

through the data collection) are Other Private and Other Public. These sources of payment categories were constructed to resolve apparent inconsistencies between individuals' reported insurance coverage and their sources of payment for specific events. For example, if the insurance variables indicated uninsured all year, but the person reported private insurance as a payer source.

2.6   File 2 Contents: Pre-imputed Expenditure Variables

Pre-imputed expenditure data are provided on File 2. Pre-imputed means that only a series of logical edits were applied to both the HC and MPC data to correct for several problems including outliers, co-payments or charges reported as total payments, and reimbursed amounts counted as out-of-pocket payments. Edits were also implemented to correct for misclassifications between Medicare and Medicaid and between Medicare HMO's and private HMO's as payment sources, as well as number of other data inconsistencies that could be resolved through logical edits. Missing data were not imputed.

The user should note that there exist only 10 sources of payment variables in the pre-imputed expenditure data, while the imputed expenditure data on File 1 contains 12 sources of payment variables. The additional two sources of payments (which are not reported as separate sources of payment through the data collection) are Other Private and Other Public. These sources of payment categories were constructed to resolve apparent inconsistencies between individuals' reported insurance coverage and their sources of payment for specific events.

The user should also note that the variable HHSFFIDX, which is the original flat fee identifier that was derived during the household interview, should be used only if user is interested in performing their own expenditure imputation.

3.0   Sample Weight (WTDPER98)

3.1   Overview

There is a single full year person-level weight (WTDPER98) 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 1998. A key person either was a member of an NHIS household at the time of the NHIS interview, or became a member of 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 Weights Construction

The final person-level weight WTDPER98 was developed in three stages. A person level weight for Panel 3 was created, including both an adjustment for nonresponse over time and poststratification, controlling to Current Population Survey (CPS) population estimates based on five variables. Variables used in the establishment of person-level poststratification control figures included: census region (Northeast, Midwest, South, West); MSA status (MSA, non-MSA); race/ethnicity (Hispanic, black

but non-Hispanic, and other); sex; and age. Then a person level weight for Panel 2 was created, again including an adjustment for nonresponse over time and poststratification, again controlling to CPS population estimates based on the same five variables. When poverty status information derived from income variables became available, a 1998 composite weight was formed from the Panel 2 and Panel 3 weights by multiplying the panel weights by .5. Then a final poststratification was done on this composite weight variable, including 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 poststratification variables in the establishment of control totals.

3.2.1   MEPS Panel 2 Weight

The person level weight for MEPS Panel 2 was developed using the 1997 full year weight for an individual as a "base" weight for survey participants present in 1997. For key, in-scope respondents who joined a RU some time in 1998 after being out-of-scope in 1997, the 1997 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 poststratification to population control figures for December 1998.   These control figures were derived by scaling back the population totals obtained from the March 1998 CPS to reflect the December, 1998 CPS estimated population distribution across age and sex categories as of December, 1998. Variables used in the establishment of person level poststratification control figures included: census region (Northeast, Midwest, South, West); MSA status (MSA, non-MSA); race/ethnicity (Hispanic, black but non-Hispanic, and other); sex, and age. Overall, the weighted population estimate for the civilian, noninstitutionalized population on December 31, 1998 is 270,114,457. Key, responding persons not in-scope on December 31, 1998 but in-scope earlier in the year retained, as their final Panel 2 weight, the weight after the nonresponse adjustment.

3.2.2   MEPS Panel 3 Weight

The person level weight for MEPS Panel 3 was developed using the MEPS Round 1 person-level weight as a 'base" weight. For key, in-scope respondents who joined a 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 1998 portion of Round 3 as well as poststratification to the same population control figures for December 1998 used for the MEPS Panel 2 weights. The same five variables employed for Panel 2 poststratification (census region, MSA status, race/ethnicity, sex, and age) were used for Panel 3 poststratification. Similarly, for Panel 3, key, responding persons not in-scope on December 31, 1998 but in-scope earlier in the year retained, as their final Panel 3 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 poststratification to figures at the family and person level obtained from the March 1998 CPS data base.

3.2.3 The Final Weight for 1998

Variables used in the establishment of person level poststratification 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, and other); sex, and age. Overall, the weighted population estimate for the civilian, noninstitutionalized population for December 31, 1998 is 270,114,457 (WTDPER98>0 and INSC1231=1). The inclusion of key, in-scope persons who were not in-scope on December 31, 1998 brings the estimated total number of persons represented by the MEPS respondents over the course of the year up to 273,533,690 (WTDPER98>0). The weighting process included poststratification to population totals obtained from the 1996 MEPS Nursing Home Component for the number of individuals admitted to nursing homes. For the 1998 full year file an additional poststratification was done to population totals obtained from the 1997 Medicare Current Beneficiary Survey (MCBS) for the number of deaths among Medicare beneficiaries experienced in the 1998 MEPS.

3.2.4 Coverage

The target population for MEPS in this file is the 1998 U.S. civilian, noninstitutionalized population. However, the MEPS sampled households are a subsample of the NHIS households interviewed in 1997 (Panel 2) and 1998 (Panel 3). New households created after the NHIS interviews for the respective panels and consisting exclusively of persons who entered the target population after 1997 (Panel 2) or after 1998 (Panel 3) are not covered by MEPS. These would include families consisting solely of: immigrants; persons leaving the military; U.S. citizens returning from residence in another country; and persons leaving institutions. It should be noted that this set of uncovered persons constitutes only a tiny proportion of the MEPS target population

4.0 Strategies for Estimation

This file is constructed for efficient estimation of utilization, expenditure, and sources of payment for home health provider visits and to allow for estimates of number of persons with home health provider visits in 1998.

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 2.5.5.2.

4.2 Basic Estimates of Utilization, Expenditure 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, expenditure and sources of payment, the value in each record contributing to the estimates must be multiplied by the weight (WTDPER98) contained on that record.

Example 1

For example, the total number of home health paid independent provider visits, for the civilian non-institutionalized population of the U.S. in 1998, is estimated as the sum of the weight (WTDPER98) across all home health paid independent provider records. That is,

$$\Sigma W_j = 4,796,013 \qquad \text{across all records with MPCELIG} = 2 \qquad (1)$$

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 (for those who had such 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,

$$(\Sigma W_j X_j)/(\Sigma W_j) = \$347.83 \qquad (2)$$

where
$$\Sigma W_j = 4,535,717 \text{ and } X_j = HHSF98X_j$$

for all home health visits by paid independent provider (MPCELIG=2) with $HHXP98X_j > 0$

This gives \$347.83 as the estimated mean amount of out-of-pocket payment of expenditures associated with home health events by paid independent providers and 4,535,717 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 non-institutionalized population of the U.S. in 1998.

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 proportion of total expenditures paid by private insurance at the home health event level. That is

$$(\Sigma\, W_j\, Y_j)/(\Sigma\, W_j) = 0.0769 \tag{3}$$

where
$$\Sigma\, W_j = 4{,}535{,}717 \quad \text{and} \quad Y_j = HHPV98X_j\,/\,HHXP98X_j$$

for all home health visits by paid independent provider (MPCELIG=2) with $HHXP98X_j > 0$

This gives 0.0769 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 non-institutionalized population of the U.S. in 1998.

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

When calculating an estimate of the total number of persons with home health events, users can use a person-level file (MEPS HC-028: Person-level Expenditures and Utilization) or the current file. However, the current 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 by paid independent providers due to a hospitalization, the current file must be used. This would be estimated as,

$$\Sigma\, W_i\, X_i \qquad \text{across all unique persons i on this file} \tag{4}$$

where
W_i is the sampling weight (WTDPER98) for person i

and

$X_i = 1$ if $HOSPITAL_j = 1$ for any home health visits by a paid independent provider (MPCELIG=2) of person i.

$ = 0$ otherwise

4.4 Person-Based Ratio Estimates

4.4.1 Person-Based Ratio Estimates Relative to Persons with Home Health Events by 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 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,

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

where

W_i is the sampling weight (WTDPER98) for person i

and

$Z_i = \sum HHXP98X_j$ across all home health visits by paid independent provider 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 MEPS File HC-028, 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 civilian non-institutionalized 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 the current file, and the denominator should be derived from data on the MEPS HC-028 person-level file. That is,

$$(\sum W_i Z_i)/(\sum W_i) \quad \text{across all unique persons i on the MEPS HC-028 file} \qquad (6)$$

where

W_i is the sampling weight (WTDPER98) for person i

and

$Z_i = 1$ if $MPCELIG_j = 2$ for any home health visits by paid independent providers 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 a 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

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 1998 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 VARSTR98 and VARPSU98, respectively. 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 VARSTR98 and VARPSU98 as the variance estimation strata and PSUs (within these strata) respectively and specifying a "with replacement" design in a computer software package SUDAAN will yield standard error estimates of $48.38 and 0.0359 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 1998 home health provider events with other 1998 MEPS public use files, including the 1998 conditions file, the 1998 prescribed medicines file, and a 1998 person-level file.

5.1 Linking a 1998 Person-Level File to the 1998 Home Health Provider Event File

Merging characteristics of interest from other 1998 MEPS files (e.g., the 1998 Full Year Population Characteristics File or the 1998 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 1998 Appendix File provides additional details on how to merge 1998 MEPS data files.

> 1. Create data set PERS by sorting a Full Year Population Characteristics File, (file HCXXX), 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 date 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  AGE  SEX  RACEX)
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 1998 Home Health Provider Event File to the 1998 Medical Conditions File and/or the 1998 Prescribed Medicines File

Due to survey design issues, there are limitations/caveats that an analyst must keep in mind when linking the different files. This limitations/caveats are listed below. For detailed linking examples, including SAS code, analysts should refer to the 1998 Appendix file.

5.3   Limitations/Caveats of RXLK (the 1998 Prescribed Medicine Link File)

The RXLK file provides a link from the 1998 prescribed medicine records to the other 1998 event files. When using RXLK, analysts should keep in mind that one home health event can link to more than one prescribed medicine record. Conversely, a prescribed medicine record may link to more than one home health event or different types of events. When this occurs, it is up to the analyst to determine how the prescribed medicine expenditures should be allocated among those medical events.

5.4   Limitations/Caveats of CLNK (the 1998 Medical Conditions Link File)

The CLNK provides a link from 1998 MEPS event files to the 1998 Medical Conditions File. When using the CLNK, 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. Users should also note that not all home health provider events link to the condition file.

References

Cohen, S.B. (1998). 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). Informing American Health Care Policy. (1999). Jossey-Bass Inc, San Francisco.

Shah, B.V., Barnwell, B.G., Bieler, G.S., Boyle, K.E., Folsom, R.E., Lavange, L., Wheeless, 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.

Attachment 1

Definitions

Dwelling Units, Reporting Units, Families, and Persons - The definitions of Dwelling Units (DUs) and Group Quarters in the MEPS Household Survey are generally consistent with the definitions employed for the National Health Interview Survey. The dwelling unit ID (DUID) is a five-digit random ID number assigned after the case was sampled for MEPS. The person number (PID) uniquely identifies all persons within the dwelling unit. The variable DUPERSID is the combination of the variables DUID and PID.

A Reporting Unit (RU) is a person or a group of persons in the sampled dwelling unit who is related by blood, marriage, adoption or other family association, and who is to be interviewed as a group in MEPS. Thus, the RU serves chiefly as a family-based "survey operations" unit rather than an analytic unit. Regardless of the legal status of their association, two persons living together as a "family" unit were treated as a single reporting unit if they chose to be so identified.

Unmarried college students under 24 years of age, who usually live in the sampled household but were living away from home and going to school at the time of the Round 1 MEPS interview, were treated as a Reporting Unit separate from that of their parents for the purpose of data collection. These variables can be found on MEPS person-level files.

In-Scope - A person was classified as in-scope (INSCOPE) if he or she was a member of the U.S. civilian, non-institutionalized population at some time during the Round 1 interview. This variable can be found on MEPS person-level files.

Keyness - The term "keyness" is related to an individual's chance of being included in MEPS. A person is key if that person is appropriately linked to the set of NHIS sampled households designated for inclusion in MEPS. Specifically, a key person either was a member of an NHIS household at the time of the NHIS interview or became a member of such a household after being out-of-scope prior to joining that household (examples of the latter situation include newborns and persons returning from military service, persons returning from an institution, or persons living outside the United States).

A non-key person is one whose chance of selection for the NHIS (and MEPS) was associated with a household that was eligible but not sampled for the NHIS, who happened to have become a member of a MEPS reporting unit by the time of the MEPS Round 1 interview. MEPS data, (e.g., utilization and income) were collected for the period of time a non-key person was part of the sampled unit to permit family level analyses. However, non-key persons who leave a sample household would not be recontacted for subsequent interviews. Non-key individuals are not part of the target sample used to obtain person-level national estimates.

It should be pointed out that a person may be key even though not part of the civilian, non-institutionalized portion of the U.S population. For example, a person in the military may be living with his or her civilian spouse and children in a household sampled for the NHIS. The person in the military would be considered a key person for MEPS. However, such a person would not receive a

person-level sample weight so long as he or she was in the military. All key persons who participated in the first round of a MEPS panel received a person-level sample weight except those who were in the military. The variable indicating "keyness" is KEYNESS. This variable can be found on MEPS person-level files.

Eligibility - The eligibility of a person for MEPS pertains to whether or not data were to be collected for that person. All key, in-scope persons of a sampled RU were eligible for data collection. The only non-key persons eligible for data collection were those who happened to be living in the same RU as one or more key persons, and their eligibility continued only for the time that they were living with a key person. The only out-of-scope persons eligible for data collection were those who were living with key in-scope persons, again only for the time they were living with a key person. Only military persons meet this description. A person was considered eligible if they were eligible at any time during Round 1. The variable indicating "eligibility" is ELIGRND1, where 1 is coded for persons eligible for data collection for at least a portion of the Round 1 reference period, and 2 is coded for persons not eligible for data collection at any time during the first round reference period. This variable can be found on MEPS person-level files.

Pre-imputed - This means that only a series of logical edits were applied to the HC data to correct for several problems including outliers, co-payments or charges reported as total payments, and reimbursed amounts counted as out-of-pocket payments. Missing data remains.

Unimputed - This means that only a series of logical edits were applied to the MPC data to correct for several problems including outliers, co-payments or charges reported as total payments, and reimbursed amounts counted as out-of-pocket payments. These data were used as the imputation source to account for missing HC data.

Imputation - A method of estimating values for cases with missing data. Hot-deck imputation creates a data set with complete data for all nonrespondent cases, by substituting the data from a respondent case that resembles the nonrespondent on certain known variables.

D. Variable-Source Crosswalk

VARIABLE-SOURCE CROSSWALK

FOR MEPS HC-026H: 1998 HOME HEALTH EVENTS PUBLIC USE FILE RELEASE

File 1:

Survey Administration Variables - Public Use

Variable	Description	Source
DUID	Dwelling unit ID (encrypted)	Assigned in sampling
PID	Person number (encrypted)	Assigned in sampling
DUPERSID	Sample person ID (DUID + PID) (encrypted)	Assigned in sampling
EVNTIDX	Event ID (encrypted)	Assigned in Sampling
EVENTRN	Event round number	CAPI derived
FFEEIDX	Flat fee ID (encrypted)	CAPI Derived

Home Health Events Variables - Public Use

Variable	Description	Source
HHBEGYR	Event start date – year	EV04/EV05
HHBEGMM	Event start date – month	EV04/EV05
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 health/home care aide	HH01
HOSPICE	Type of health care worker – hospice worker	HH01

Variable	Description	Source
HMEMAKER	Type of health care worker- homemaker	HH01
IVTHP	Type of health care worker – IV or infusion therapist	HH01
MEDLDOC	Type of health care worker – medical doctor	HH01
NURPRACT	Type of health care worker – nurse/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 therapist	HH01
RESPTHP	Type of health care worker – respiratory therapist	HH01
SOCIALW	Type of health care worker – social worker	HH01
SPEECTHP	Type of health care worker – speech therapist	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	HH02
OTHCW	Type of health care worker – some other type of health care worker	HH02
OTHCWOS	Specify other type of health care worker	HH02
HOSPITAL	Any home health care provider event due to hospitalization	HH03
VSTRELCN	Any home health care provider event related to a health condition	HH04
TREATMT	Person received medical treatment	HH06
MEDEQUIP	Person was taught how to use medical equipment	HH07
DAILYACT	Person was helped with daily activities	HH08
COMPANY	Person received companionship services	HH09
OTHSVCE	Person received other home health care services	HH10
OTHSVCOS	Specify other home health care service received	HH10
FREQCY	Provider helped person every week/some weeks	HH11

Variable	Description	Source
DAYSPWK	Number of days per week provider came (agency events only)	HH12
DAYSPMO	Number of days per month provider came (agency events only)	HH13
HOWOFTEN	Provider came once per day or more than once per day	HH14
TMSPDAY	Times per day provider came to home to help	HH15
HRSLONG	Hours each visit lasted	HH16
MINLONG	Minutes each visit lasted	HH16
SAMESVCE	Any other months person received services	HH17
HHDAYS	Number of days person received care per month for that event	Constructed

Imputed Expenditure Variables – Public Use

Variable	Description	Source
FFHHTYPE	Flat fee bundle - stem or leaf indicator (edited)	FF01 or FF02 (edited)
FFBEF98	Total number of visits in flat fee before 1998	FF05
FFTOT99	Total # of visits in flat fee after 1998	FF02 (edited)
HHSF98X	Amount paid, family *note: rounded to cents*	CP11 (Edited/Imputed)
HHMR98X	Amount paid, Medicare *note: rounded to cents*	CP09 (Edited/Imputed)
HHMD98X	Amount paid, Medicaid *note: rounded to cents*	CP07 (Edited/Imputed)
HHPV98X	Amount paid, private insurance *note: rounded to cents*	CP07 (Edited/Imputed)
HHVA98X	Amount paid, Veterans *note: rounded to cents*	CP07 (Edited/Imputed)
HHCH98X	Amount paid, CHAMPUS/CHAMPVA *note: rounded to cents*	CP07 (Edited/Imputed)
HHOF98X	Amount paid, other federal *note: rounded to cents*	CP07 (Edited/Imputed)
HHSL98X	Amount paid, state and local government *Note: rounded to cents*	CP07 (Edited/Imputed)
HHWC98X	Amount paid, worker's compensation *Note: rounded to cents*	CP07 (Edited/Imputed)
HHOR98X	Amount paid, other private *Note: rounded to cents*	Constructed
HHOU98X	Amount paid, other public *Note: rounded to cents*	Constructed
HHOT98X	Amount paid, other insurance *Note: rounded to cents*	CP07 (Edited/Imputed)
HHXP98X	Sum of payments HHSF98X – HHOT98X *Note: rounded to cents*	Constructed
HHTC98X	Total charge for visit *Note: rounded to cents*	CP09 (Edited/Imputed)

Weights - Public Use

Variable	Description	Source
WTDPER98	Person weight full-year 1998 (poverty/mortality adjusted)	Constructed
VARPSU98	Variance estimation PSU 1998	Constructed
VARSTR98	Variance estimation stratum, 1998	Constructed

File 2:

Survey Administration Variables - Public Use

Variable	Description	Source
DUID	Dwelling unit ID (encrypted)	Assigned in sampling
PID	Person number (encrypted)	Assigned in sampling
DUPERSID	Sample person ID (DUID + PID) (encrypted)	Assigned in sampling
EVNTIDX	Event ID (encrypted)	Assigned in Sampling
HHSFFIDX	Household reported flat fee id (unedited) (encrypted)	CAPI Derived

Pre-imputed Expenditure Variables

Variable	Description	Source
HHSF98H	Amount paid, family (pre-imputed) *note: rounded to cents*	CP11 (Edited)
HHMR98H	Amount paid, Medicare (pre-imputed) *note: rounded to cents*	CP09 (Edited)
HHMD98H	Amount paid, Medicaid (pre-imputed) *note: rounded to cents*	CP07 (Edited)
HHPV98H	Amount paid, private insurance (pre-imputed) *note: rounded to cents*	CP07 (Edited)
HHVA98H	Amount paid, Veterans (pre-imputed) *note: rounded to cents*	CP07 (Edited)
HHCH98H	Amount paid, CHAMPUS/CHAMPVA (pre-imputed) *note: rounded to cents*	CP07 (Edited)
HHOF98H	Amount paid, other federal (pre-imputed) *note: rounded to cents*	CP07 (Edited)
HHSL98H	Amount paid, state and local government (pre-imputed) *note: rounded to cents*	CP07 (Edited)
HHWC98H	Amount paid, worker's compensation (pre-imputed) *note: rounded to cents*	CP07 (Edited)
HHOT98H	Amount paid, other insurance (pre-imputed) *note: rounded to cents*	CP07 (Edited)
HHUC98H	Amount paid, uncollected liability (pre-imputed)	CP07 (Edited)
HHTC98H	Total charge (pre-imputed) *note: rounded to cents*	CP09 (Edited)

Weights – Public Use

Variable	Description	Source
WTDPER98	Person weight full-year 1998 (poverty/mortality adjusted)	Constructed
VARSTR98	Variance estimation stratum, 1998	Constructed
VARPSU98	Variance estimation PSU 1998	Constructed