

**MEPS HC-033G:
1999 Office-Based Medical Provider Visits
July 2002**

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

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-3
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	Codebook Structure	C-3
2.2	Reserved Codes.....	C-3
2.3	Codebook Format	C-4
2.4	Variable Naming.....	C-4
2.4.1	General.....	C-4
2.4.2	Expenditure and Sources of Payment Variables.....	C-5
2.5	File Contents.....	C-6
2.5.1	Survey Administration and ID Variables.....	C-6
2.5.1.1	Person Identifiers (DUID- DUPERSID).....	C-6
2.5.1.2	Record Identifiers (EVNTIDX- FFEEDIX)	C-6
2.5.1.3	Record Indicators (EVENTRN – OBR2FLAG)	C-6
2.5.2	Characteristics of Office-Based Medical Provider Visits.....	C-7
2.5.2.1	Date of Office-Based Medical Provider Visit (OBDATEYR - OBDATEDD)	C-7
2.5.2.2	Visit Details (SEETLKPV-VSTRELCN).....	C-7
2.5.2.3	Treatments, Procedures, Services, and Prescription Medicines (PHYSTH-MEDPRESC).....	C-7
2.5.2.4	Other Visit Details (VAPLACE).....	C-7
2.5.2.5	MPC Indicator (MPCELIG, MPCDATA).....	C-8
2.5.3	Condition and Procedure Codes (OBICD1X-OBICD4X, OBPRO1X) and Clinical Classification Codes (OBCCC1X-OBCCC4X)	C-8
2.5.4	Flat Fee Variables (FFOBTYP, FFBEF99, FFTOT99).....	C-9
2.5.4.1	Definition of Flat Fee Payments.....	C-9
2.5.4.2	Flat Fee Variable Descriptions.....	C-9
2.5.4.3	Caveats of Flat Fee Groups.....	C-10
2.5.5	Expenditure Data.....	C-10
2.5.5.1	Definition of Expenditures	C-10
2.5.5.2	Imputation and Data Editing Methodologies of Expenditure Variables	C-11
2.5.5.2.1	General Data Editing Methodology	C-11
2.5.5.2.2	General Hot-Deck Imputation.....	C-11
2.5.5.3	Capitation Imputation	C-12
2.5.5.4	Imputation Flag Variable (IMPFLAG)	C-12

2.5.5.5	Flat Fee Expenditures	C-12
2.5.5.6	Zero Expenditures	C-12
2.5.5.7	Discount Adjustment Factor.....	C-13
2.5.5.8	Sources of Payment	C-13
2.5.5.9	Office- Based Expenditure Variables (OBSF99X - OBXP99X).....	C-14
2.5.5.10	Rounding.....	C-14
3.0	Sample Weight (PERWT99F).....	C-15
3.1	Overview.....	C-15
3.2	Details on Person Weights Construction	C-15
3.2.1	MEPS Panel 3 Weight	C-15
3.2.2	MEPS Panel 4 Weight	C-16
3.2.3	The Final Weight for 1999	C-16
3.2.4	Coverage	C-16
4.0	Strategies for Estimation	C-17
4.1	Variables with Missing Values	C-17
4.2	Basic Estimates of Utilization, Expenditure and Sources of Payment.....	C-17
4.3	Estimates of the Number of Persons with Office-Based Medical Provider Visits.....	C-18
4.4	Person-Based Ratio Estimates	C-19
4.4.1	Person-Based Ratio Estimates Relative to Persons with Office-Based Medical Provider Visits.....	C-19
4.4.2	Person-Based Ratio Estimates Relative to the Entire Population.....	C-19
4.5	Sampling Weights for Merging Previous Releases of MEPS Household Data with this Event File.....	C-19
4.6	Variance Estimation	C-20
5.0	Merging/Linking MEPS Data Files	C-21
5.1	Linking a Person-Level File to the Office-Based Medical Provider Visit File	C-21
5.2	Linking the Office-Based Medical Provider Visit file to the MEPS 1999 Medical Conditions File and/or the MEPS 1999 Prescribed Medicines File.....	C-22
5.3	Limitations/Caveats of RXLK (the Prescribed Medicine Link File).....	C-22
5.4	Limitations/Caveats of CLNK (the Medical Conditions Link File).....	C-22
References	C-23
Attachment 1	C-25
D. Variable Source Crosswalk.....		D-1

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 an 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. 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 2 1/2 -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.

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. Electronic files and accompanying documentation are available from the Internet on the MEPS web site: <<http://www.meps.ahrq.gov/>>.

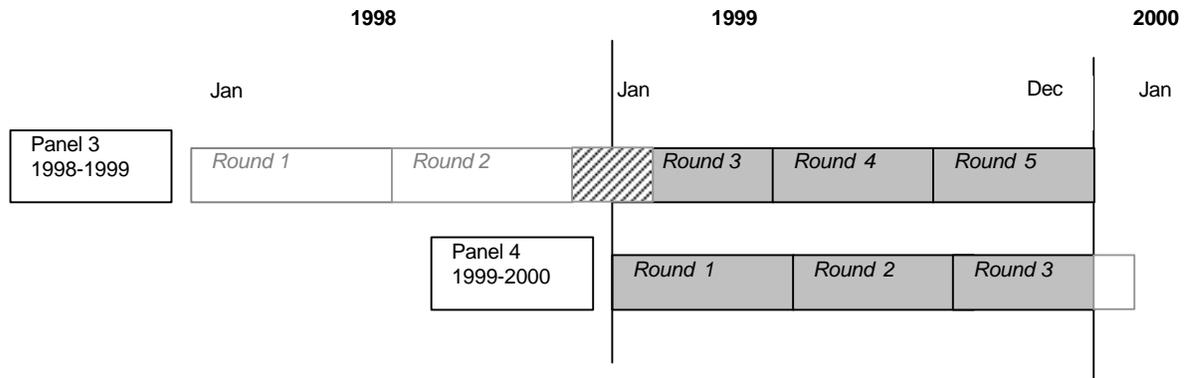
Additional information on MEPS is available from the MEPS project manager or the MEPS public use data manager at the Center for Cost and Financing Studies, Agency for Healthcare Research and Quality.

C. Technical and Programming Information

1.0 General Information

This documentation describes one in a series of public use event files from the 1999 Medical Expenditure Panel Survey (MEPS) Household (HC) and Medical Provider Components (MPC). Released as an ASCII data file and a SAS transport file, the 1999 Office-Based Medical Provider public use event file provides detailed information on office-based provider visits for a nationally representative sample of the civilian noninstitutionalized population of the United States. Data from the office-based provider events file can be used to make estimates of office-based provider utilization and expenditures for calendar year 1999.

As illustrated below, this file consists of MEPS survey data obtained in the 1999 portion of Round 3 (Round 2 for some cases, see OBR2FLAG) and Rounds 4 and 5 for Panel 3, as well as Rounds 1, 2 and the 1999 portion of Round 3 for Panel 4 (i.e., the rounds for the MEPS panels covering calendar year 1999).



 **NOTE:** Typically for MEPS panels, MEPS Round 2 data collection ends in the first year of a panel and Round 3 data collection begins in the first year of the panel and crosses the year boundary into the second year of the panel. The crosshatched area in the above figure signifies that Round 2 data collection for approximately one quarter of the Panel 3 households began in 1998, the first year of the panel, but ended in 1999. For those households, all of the Round 3 data collection occurred in 1999. For the other three quarters of Panel 3 households, Round 2 data collection followed the typical pattern and began and ended in 1998. For those households, Panel 3 Round 3 data collection took place during both the first and second years of the panel, as is typically done for Round 3.

Each record on this event file represents a unique office-based provider event; that is, an office-based provider event reported by the household respondent. Counts utilization of office-based provider visits are based entirely on household reports. Office-based 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 the MPC. Information from the MPC was used to supplement expenditure payment data, on the office-based provider file, reported by the household.

Data from this event file can be merged with other 1999 MEPS HC data files, for purposes of appending person-level data such as demographic characteristics or health insurance coverage to each office-based provider visit record on the current file.

This file can be also used to construct summary variables of expenditures, sources of payment, and related aspects of office-based provider visits for calendar year 1999. Aggregate annual person-level information on the use of office-based providers and other health services use is provided on the MEPS 1999 Full Year Person Level Expenditure file, where each record represents a MEPS sampled person.

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

- Data File Information
- Sample Weights and Variance Estimation Variables
- Strategies for Estimation
- Merging/linking MEPS Data Files
- References
- Attachment 1: Definitions
- Codebooks
- Variable to Source Crosswalk

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

2.0 Data File Information

The office-based provider public use data set contains 100,312 office-based provider event records; of these records, 98,154 are associated with persons having a positive person-level weight (PERWT99F). This file includes office-based provider event records for all household survey respondents who resided in eligible responding households and reported at least one office-based provider event. Each record represents one household-reported office-based provider event that occurred during calendar year 1999. Office-based provider visits known to have occurred after December 31, 1999 are not included on this file. Some household respondents may have multiple events and thus will be represented in multiple records on this file. Other household respondents may have reported no events and thus will have no records on this file. These data were collected during the 1999 portion of Round 3 (Round 2 for some cases, see OBR2FLAG), and Rounds 4 and 5 for Panel 3, as well as Rounds 1, 2, and the 1999 portion of Round 3 for Panel 4 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 1999 eligibility (i.e., persons with a positive 1999 full-year person-level sampling weight (PERWT99F>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 1999 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 1999 eligibility, and at least one family member had a positive 1999 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 1999 full-year family-level weight (WTFAM99>0)).

Please refer to Attachment 1 for definitions of keyness, in-scope and eligibility.

Each office-based medical provider event record includes the following: date of the event; type of provider seen; time spent with the provider; type of care received; types of treatments (i.e. physical therapy, occupational therapy, speech therapy, chemotherapy, radiation therapy etc.) received during the event; type of services (i.e., lab test, sonogram or ultrasound, x-rays etc) received, medicines prescribed during the event; flat fee information, imputed sources of payment, total payment and total charge of the office-based event expenditure; and a full-year person-level weight.

Data from the HC-033G file can be merged with the MEPS 1999 Full Year Population Characteristics file using the unique person identifier, DUPERSID, to append person-level characteristics such as demographic or health insurance characteristics to each record. The office-based medical provider events can also be linked to the MEPS 1999 Medical Conditions File and MEPS 1999 Prescribed Medicines File. Please see the section 5.0 for details on how to merge MEPS data files.

Panel 3 cases (PANEL99 = 3 on the MEPS 1999 Full Year Population Characteristics File) can also be linked back to the 1998 MEPS HC public use data files. However, data users/analysts should be aware that, at this time, no weight is being provided to facilitate two-year analysis of Panel 3 data.

2.1 Codebook Structure

For each variable on the office-based provider file, both weighted and unweighted frequencies are provided in the codebook. The codebook and data file sequence list variables in the following order:

- Unique person identifiers
- Unique office-based medical provider event identifier
- Other survey administration variables
- Office-based medical provider characteristic variables
- ICD-9 codes
- Clinical Classification Software codes
- 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 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/>>).

2.3 Codebook Format

The office-based medical provider 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 this file were derived from the HC survey questionnaire. The source of each variable is identified in Section D, the “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 as “capi derived” or “assigned in sampling,”;
- 2) variables which come from one or more specific questions have those questionnaire sections and question numbers indicated in the “Source” column
- 3) EV-Event Roster section
- 4) FF- Flat Fee section
- 5) CP- Charge Payment section;

- 6) variables constructed from multiple questions using complex algorithms are labeled “Constructed” in the “Source” column; and
- 7) (4) variables which have been edited or imputed are so indicated.

2.4.2 Expenditure and Sources of Payment Variables

Imputed versions of the expenditure and sources of payment variables are provided on this file. These variable names follow a standard naming convention and are 8 characters in length. All imputed variables end with an “X” indicating they are fully edited and imputed.

The total sum of payments variables, 12 sources of payment variables, and the 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 visit	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	XP - sum of payments
MR - Medicare	SL - State/local government	
MD - Medicaid	WC – Worker’s Compensation	
PV - private insurance	OT - other insurance	
VA - Veterans	OR - other private	
CH - CHAMPUS/CHAMPVA	OU - other public	

The fifth and sixth characters indicate the year (99). The last character indicates whether it is edited/imputed (X).

For example, OBSF99X is the edited/imputed amount paid by self or family for an office-based medical provider expenditure incurred in 1999.

2.5 File Contents

2.5.1 Survey Administration and ID Variables

2.5.1.1 Person Identifiers (DUID- 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 the documentation for the 1999 Full Year Population Characteristics File or to definitions listed in Attachment 1.

2.5.1.2 Record Identifiers (EVNTIDX- FFEEIDX)

EVNTIDX uniquely identifies each office-based medical provider event (i.e. each record on the office-based medical provider file) and is the variable required to linking office-based medical provider events to data files containing details on conditions and/or prescribed medicines (MEPS 1999 Medical Condition file and MEPS 1999 Prescribed Medicine file; respectively). For details on linking see Section 5.0 or the MEPS 1999 Appendix file.

FFEEIDX is a constructed variable which 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 events (the prenatal visit, the delivery, and the postpartum visits) would have the same value for FFEEIDX. FFEEIDX identifies a flat fee payment situation that was identified using information from the Household Component. Please note that FFEEIDX should be used to link up all MEPS event files (excluding prescribed medicines) in order to determine the full set of events that are part of a flat fee group.

2.5.1.3 Record Indicators (EVENTRN – OBR2FLAG)

EVENTRN indicates the round in which the outpatient event was first reported. Please note: Rounds 3 (Round 2 for some cases, see OBR2FLAG), 4, and 5 are associated with MEPS survey data collected from Panel 3. Likewise, Round 1, 2, and 3 are associated with data collected from Panel 4.

OBR2FLAG indicates whether or not a Panel 3 Round 2 event occurred in 1999. OBR2FLAG was assigned a value =1 where an event in Round 2 of Panel 3 occurred in a portion of calendar year 1999. Events from Panel 4 will also have OBR2FLAG = -1. Typically, only Round 3 of a MEPS panel covers two calendar years, so the OBR2FLAG was developed to identify where data collection procedures were modified. All utilization data for calendar year 1999 is provided on this file regardless of the round in which it happened to be collected. Data users/analysts need not modify any procedures to deal with this departure from the usual data collection process as the event variables have been developed so that the process is transparent.

2.5.2 Characteristics of Office-Based Medical Provider Visits

2.5.2.1 Date of Office-Based Medical Provider Visit (OBDATEYR - OBDATEDD)

The file contains variables describing office-based medical provider events reported by respondents in the Medical Provider Visits section of the MEPS HC survey questionnaire. There are three variables which indicate the day, month and year an office-based provider visit occurred (OBDATEYR, OBDATEMM, and OBDATEDD, respectively). These variables have not been edited or imputed.

2.5.2.2 Visit Details (SEETLKPV-VSTRELCN)

The questionnaire determines if during the office-based medical provider visit whether the person actually saw the provider or talked to the provider on the telephone (SEETLKPV). It also establishes if the person was referred by another physician or medical provider (REFERDBY), and whether the person saw or spoke to a medical doctor or not (SEEDOC). If the person did not see a physician (i.e., a medical doctor), the respondent was asked to identify the type of medical person seen (MEDPTYPE). The respondent was also asked how much time was spent with the medical provider (TIMESPNT). Whether or not any medical doctors worked at the visit location (DOCATLOC), the type of care the person received (VSTCTGRY), and whether or not the visit or telephone call was related to a specific condition (VSTRELCN) were also determined.

2.5.2.3 Treatments, Procedures, Services, and Prescription Medicines (PHYSTH-MEDPRESC)

Types of treatments received during the office-based medical provider visit include physical therapy (PHYSTH), occupational therapy (OCCUPTH), speech therapy (SPEECHTH), chemotherapy (CHEMOTH), radiation therapy (RADIATTH), kidney dialysis (KIDNEYD), IV therapy (IVTHER), drug or alcohol treatment (DRUGTRT), allergy shots (RCVSHOT), and psychotherapy/counseling (PSYCHOTH). Services received during the visit included whether or not the person received lab tests (LABTEST), a sonogram or ultrasound (SONOGRAM), x-rays (XRAYS), a mammogram (MAMMOG), an MRI or a CAT scan (MRI), an electrocardiogram (EKG), an electroencephalogram (EEG), a vaccination (RCVVAC), anesthesia (ANESTH), or other diagnostic tests or exams (OTHSVCE). Minimal editing was done across treatment, services, and procedures to ensure consistency across inapplicables, not ascertained, don't know, refused, and no services received values. Whether or not a surgical procedure was performed during the visit was asked (SURGPROC) and, if so, the procedure name (SURGNAME). Finally, the questionnaire determined if a medicine was prescribed for the person during the visit (MEDPRESC).

2.5.2.4 Other Visit Details (VAPLACE)

VAPLACE is a constructed variable that indicates whether the provider worked at a VA facility. This variable only has valid data for providers that were sampled into the Medical Provider Component. All other providers are classified as unknown.

2.5.2.5 MPC Indicator (MPCELIG, MPCDATA)

MPCELIG is a constructed variable that indicates whether the office-based provider visit was eligible for MPC data collection. MPCDATA is a constructed variable that indicates whether or not MPC data was collected for the office-based provider.

2.5.3 Condition and Procedure Codes (OBICD1X-OBICD4X, OBPRO1X) and Clinical Classification Codes (OBCCC1X-OBCCC4X)

Information on household reported medical conditions and procedures associated with each office-based medical provider visit are provided on this file. There are up to four condition codes (OBICD1X-OBICD4X), one procedure code (OBPRO1X), and up to four clinical classification codes (OBCCC1X-OBCCC4X) listed for each office-based medical provider visit. In order to obtain complete condition information associated with an event, the analyst must link to the Medical Conditions File. Details on how to link to the MEPS Medical Conditions File are provided in section 5.0. The user should note that due to confidentiality restrictions, provider reported condition information is not publicly available.

The medical conditions reported by the Household Component respondent were recorded by the interviewer as verbatim text, which were then coded to fully-specified 1999 ICD-9-CM codes, including medical condition and V codes (see Health Care Financing Administration, 1980), by professional coders. Although codes were verified and error rates did not exceed 2.5 percent for any coder, data users/analysts should not presume this level of precision in the data; the ability of household respondents to report condition data that can be coded accurately should not be assumed (see Cox and Cohen, 1985; Cox and Iachan, 1987; Edwards, et al, 1994; and Johnson and Sanchez, 1993). For detailed information on conditions, please refer to the documentation on the 1999 Medical Conditions File. For frequencies of conditions by event type, please see: the MEPS 1999 Appendix File.

The ICD-9-CM codes were aggregated into clinically meaningful categories. These categories, included on the file as OBCCC1X-OBCCC4X, were generated using Clinical Classification Software (formerly known as Clinical Classifications for Health Care Policy Research (CCHPR)), (Elixhauser, et al., 1998), which aggregates conditions and V-codes into 260 mutually exclusive categories, most of which are clinically homogeneous.

In order to preserve respondent confidentiality, nearly all of the condition codes provided on this file have been collapsed from fully-specified codes to 3-digit code categories. The reported ICD-9-CM code values were mapped to the appropriate clinical classification category prior to being collapsed to the 3-digit categories.

The condition codes (and clinical classification codes) and procedure codes linked to each office-based medical provider visit event are sequenced in the order in which the conditions were reported by the household respondent, which was in chronological order of occurrence and not in order of importance or severity. Data user/analysts who use the Medical Conditions file in conjunction with

this office-based medical provider visit file should note that the order of conditions on this file is not identical to that on the 1999 Medical Conditions file.

2.5.4 Flat Fee Variables (FFOBTYP, FFBEF99, FFTOT99)

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 services provided during a defined period of time. Examples would be an 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 the office-based provider file, includes flat fee groups where at least one of the health care events, as reported by the HC respondent, occurred during 1999. By definition, a flat fee group can span multiple years and/or event types (e.g., outpatient department visit, physician office visit). Furthermore a single person can have multiple flat fee groups.

There are four variables on the office-based provider file that describe a flat fee payment situation and the number of medical events that are part of a flat fee group.

2.5.4.2 Flat Fee Variable Descriptions

Flat Fee ID (FFEEIDX)

As noted earlier in the Section 2.5.1.2 "Record Identifiers," for a person, the variable FFEEIDX can be used to uniquely identify all events that are part of the same flat fee group. It can identify such events from all of the 1999 MEPS event files (excluding the prescribed medicine file) because FFEEIDX is the same value on all of the MEPS event files. For the office-based medical provider events that are not part of a flat fee payment situation, the flat fee variables described below are all set to -1 INAPPLICABLE.

Flat Fee Type (FFOBTYP)

FFOBTYP indicates whether the 1999 office-based medical provider event is the "stem" or "leaf" of a flat fee group. A stem (records with FFOBTYP = 1) is the initial medical service (event) which is followed by other medical events that are covered under the same flat fee payment. The leaves of the flat fee group (records with FFOBTYP = 2) are those medical events that are tied back to the initial medical event (the stem) in the flat fee group. These "leaf" records have their expenditure variables set to zero.

Counts of Flat Fee Events that Cross Years (FFBEF99 - FFTOT99)

As described in Section 2.5.4.1, a flat fee payment situation covers multiple events and the multiple events could span multiple years. For situations where a 1999 office-based medical provider visit is part of a group of events, and some of the events occurred before 1999, counts of the known events

are provided on the office-based medical provider event file record. Indicator variables are provided if some of the events occurred before or after 1999. These variables are:

FFBEF99 -- total number of pre-1999 events in the same flat fee group as the 1999 office-based medical provider event. This count would not include 1999 office-based medical provider visit.

FFTOT00 -- indicates whether or not there are 2000 medical events in the same flat fee group as the 1999 office-based medical provider event record.

2.5.4.3 Caveats of Flat Fee Groups

Data users/analysts should note that flat fee payment situations are common on the office-based medical provider file. There are 2,534 office-based medical provider events that are identified as being part of a flat fee payment group. In order to correctly identify all events that are part of a flat fee group, the user should link all MEPS events, except the prescribed medicine file, using the variable FFEEIDX.

In general, every flat fee group should have an initial visit (stem) and at least one subsequent visit (leaf). There are some situations where this is not true. For some of these flat fee groups, the initial visit reported occurred in 1999, but the remaining visits that were part of this flat fee group occurred in 2000. In this case, the 1999 flat fee group represented on this file would consist of one event (the stem). The 2000 “leaf” 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 visit began in 1998 but subsequent visits occurred during 1999. In this case, the initial visit would not be represented on the file. This 1999 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. In a small number of cases, there are flat fee groups that span various event types (e.g. office-based medical provider visit). The stem may have been reported as one event type (e.g. outpatient department visit) and the leaves may have been reported as another event type (e.g. office-based medical provider visit).

2.5.5 Expenditure Data

2.5.5.1 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, the estimates do not incorporate any payment not directly tied to specific medical care visits, 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 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). AHRQ has developed factors to apply to the 1987 NMES expenditure data to facilitate longitudinal analysis. These factors can be assessed via CCFS data center. For more information see the data center section of the MEPS web site <<http://www.meps.ahrq.gov>>.

2.5.5.2 Imputation and Data Editing Methodologies of Expenditure Variables

The expenditure data included on this file were derived from both the MEPS HC and MPC. The MPC contacted medical providers identified by household respondents. The charge and payment data from medical providers was used in the expenditure imputation process to supplement missing household data. For all office-based medical provider visits, MPC data were used if complete; otherwise HC data were used if complete. Missing data for office-based medical provider visits where HC data were not complete and MPC data were not collected or complete were derived through the imputation process. Specific methodologies for editing and imputing office-based provider expenditures follows.

2.5.5.2.1 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 HMOs and private HMOs 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.

2.5.5.2.2 General Hot-Deck Imputation

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. 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. 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.

2.5.5.3 Capitation Imputation

The imputation process was also used to make expenditure estimates at the event level for events that were paid on a capitated basis. The capitation imputation procedure was designed as a reasonable approach to complete event level expenditures for respondents in managed care plans. The procedure was conducted in two stages. First, HMO events reported in the MPC as covered by capitated arrangements were imputed using similar MPC HMO events that were paid on a fee-for-service basis, with total charge as a key variable. Then, this completed set of MPC events was used as the donor pool for unmatched household-reported events for sample persons in HMOs. By using this strategy, capitated HMO events were imputed as if the provider were reimbursed from the HMO on a discounted fee-for-service basis.

2.5.5.4 Imputation Flag Variable (IMPFLAG)

Unlike prior data releases, only one imputation flag was created for 1999 event files. This variable, IMPFLAG, is a six-category variable that indicates if the event contains complete Household Component (HHC) or Medical Provider Component (MPC) data, was fully or partially imputed, or was imputed in the capitated imputation process. Following is how the new imputation flag is coded; the categories are mutually exclusive.

IMPFLAG= 0 (not eligible for imputation)
IMPFLAG=1 (complete HC data)
IMPFLAG=2 (complete MPC data)
IMPFLAG=3 (fully imputed)
IMPFLAG=4 (partially imputed)
IMPFLAG=5 (capitation imputation)

2.5.5.5 Flat Fee Expenditures

The approach used to count expenditures for flat fees was to place the expenditure on the first visit of the flat fee group. The remaining visits have zero payments. Thus, if the first visit in the flat fee group occurred prior to 1999, all of the events that occurred in 1999 will have zero payments. Conversely, if the first event in the flat fee group occurred at the end of 1999, the total expenditure for the entire flat fee group will be on that event, regardless of the number of events it covered after 1999.

2.5.5.6 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 visits 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.

2.5.5.7 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.

2.5.5.8 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) 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 Source - 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, 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 persons who were not enrolled in Medicaid, but were presumed eligible by a provider who ultimately received payments from the program.

2.5.5.9 Office- Based Expenditure Variables (OBSF99X - OBP99X)

There are 13 expenditure variables included on this event file. 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 (OBXP99X) which for each office-based medical provider visit sums all the expenditures from the various source of payment. The 12 sources of payment expenditure variables for each office-based medical provider visit are the following: amount paid by self or family (OBSF99X), amount paid by Medicare (OBMR99X), amount paid by Medicaid (OBMD99X), amount paid by private insurance (OBPV99X), amount paid by Veterans Administration (OBVA99X), amount paid by CHAMPUS/CHAMPVA (OBCH99X), amount paid other federal sources (OBOF99X), amount paid by state and local (non-federal) government sources (OBSL99X), amount paid by Worker’s Compensation (OBWC99X), and amount paid by some other source of insurance (OBOT99X). As mentioned previously, there are two additional expenditure variables called OBOR99X and OBOU99X (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.

2.5.5.10 Rounding

Expenditure variables have been rounded to the nearest penny. Person-level expenditure information released on the MEPS 1999 Person Level Expenditure file will be rounded to the nearest dollar. It should be noted that using the MEPS event files to create person-level totals will yield slightly different totals than that those found on 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 1999 Appendix File for details on such rounding differences.

3.0 Sample Weight (PERWT99F)

3.1 Overview

There is a single full year person-level weight (PERWT99F) 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 1999. 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 person-level weight PERWT99F was developed in three stages. A person level weight for Panel 4 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 3 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 1999 composite weight was formed from the Panel 3 and Panel 4 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 3 Weight

The person level weight for MEPS Panel 3 was developed using the 1998 full year weight for an individual as a “base” weight for survey participants present in 1998. For key, in-scope respondents who joined a RU some time in 1999 after being out of scope in 1998, the 1998 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 1999. These control figures were derived by scaling back the population totals obtained from the March 1999 CPS to reflect the December, 1999 CPS estimated population distribution across age and sex categories as of December, 1999. 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, 1999 is 273,003,778. Key, responding persons not in-scope on December 31, 1999 but in-scope earlier in the year retained, as their final Panel 3 weight, the weight after the nonresponse adjustment.

3.2.2 MEPS Panel 4 Weight

The person level weight for MEPS Panel 4 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 1999 portion of Round 3 as well as poststratification to the same population control figures for December 1999 used for the MEPS Panel 3 weights. The same five variables employed for Panel 3 poststratification (census region, MSA status, race/ethnicity, sex, and age) were used for Panel 4 poststratification. Similarly, for Panel 4, key, responding persons not in-scope on December 31, 1999 but in-scope earlier in the year retained, as their final Panel 4 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 1999 CPS data base.

3.2.3 The Final Weight for 1999

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, 1999 is 273,003,778 (PERWT99F>0 and INSC1231=1). The inclusion of key, in-scope persons who were not in-scope on December 31, 1999 brings the estimated total number of persons represented by the MEPS respondents over the course of the year up to 276,410,767 (PERWT99F>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 1999 full year file an additional poststratification was done to population totals obtained from the 1998 Medicare Current Beneficiary Survey (MCBS) for the number of deaths among Medicare beneficiaries experienced in the 1999 MEPS.

3.2.4 Coverage

The target population for MEPS in this file is the 1999 U.S. civilian, noninstitutionalized population. However, the MEPS sampled households are a subsample of the NHIS households interviewed in 1998 (Panel 3) and 1999 (Panel 4). New households created after the NHIS interviews for the respective Panels and consisting exclusively of persons who entered the target population after 1998 (Panel 3) or after 1999 (Panel 4) 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 office-based medical provider visits and to allow for estimates of number of persons with office-based medical provider visits in 1999.

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.

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 office-based medical provider visits utilization, expenditure and sources of payment, the value in each record contributing to the estimates must be multiplied by the weight (PERWT99F) contained on that record.

Example 1

For example, the total number of office-based medical provider visits, for the civilian non-institutionalized population of the U.S. in 1999, is estimated as the sum of the weight (PERWT99F) across all office-based medical provider records. That is,

$$\sum W_j = 1,237,896,836 \quad (1)$$

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 office-based medical provider visit (for those who had such expense greater than 0) should be calculated as the weighted mean of the office-based provider's bill paid by self/family. That is,

$$(\sum W_j X_j)/(\sum W_j) = \$19.86 \quad (2)$$

where

$$\sum W_j = 1,152,705,261 \quad \text{and} \quad X_j = \text{OBSF99}X_j$$

for all records with $\text{OBXP99}X_j > 0$

This gives \$19.86 as the estimated mean amount of out-of-pocket payment of expenditures associated with office-based medical provider visits and 1,152,705,261 as an estimate of the total number of office-based medical provider visits with expenditure. Both of these estimates are for the civilian non-institutionalized population of the U.S. in 1999.

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 office-based medical provider visits. This should be calculated as the weighted mean of proportion of total expenditures paid by private insurance at the provider visit level. That is

$$(\sum W_j Y_j)/(\sum W_j) = 0.4380 \quad (3)$$

where

$$\sum W_j = 1,152,705,261 \quad \text{and} \quad Y_j = \text{OBPV99}X_j / \text{OBXP99}X_j$$

for all office-based medical provider visits with $\text{OBXP99}X_j > 0$

This gives 0.4380 as the estimated mean proportion of total expenditures paid by private insurance for office-based medical provider visits with expenditures for the civilian non-institutionalized population of the U.S. in 1999.

4.3 Estimates of the Number of Persons with Office-Based Medical Provider Visits

When calculating an estimate of the total number of persons with office-based medical provider visits, users can use a person-level 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 office-based medical provider visits in person and not by telephone, the current 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 (PERWT99F) for person i

and

$X_i = 1$ if $\text{SEETLKP}V_j = 1$ for any office-based medical provider visit of person i .

= 0 otherwise

4.4 Person-Based Ratio Estimates

4.4.1 Person-Based Ratio Estimates Relative to Persons with Office-Based Medical Provider Visits

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 office-based medical provider visits is estimated as,

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

where

W_i is the sampling weight (PERWT99F) for person i

and

$Z_i = \sum \text{OBXP99X}_j$ across all office-based medical provider visits 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 office-based medical provider visit are represented on this data file. In this case a person level file, which has data for all sampled persons, must be used to estimate the total number of persons (i.e. those with visits and those without visits). For example, to estimate the proportion of civilian non-institutionalized population of the U.S. with at least one in person office-based medical provider visit, the numerator would be derived from data on the current file, and the denominator should be derived from data on the person level file. That is,

$$(\sum W_i Z_i)/(\sum W_i) \quad \text{across all unique persons } i \text{ on the person level file} \quad (6)$$

where

W_i is the sampling weight (PERWT99F) for person i

and

$Z_i = 1$ if SEETLKPV_j = 1 for any office-based medical provider visit 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 1999 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 VARSTR99 and VARPSU99, 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 VARSTR99 and VARPSU99 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 \$0.5767 and 0.0086 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 office-based medical provider file can be used alone or in conjunction with other files. This section provides instructions for linking the office-based medical provider visits with other MEPS public use files, including the conditions file, the prescribed medicines file, and a person-level file.

5.1 Linking a Person-Level File to the Office-Based Medical Provider Visit File

Merging characteristics of interest from other MEPS files (e.g., 1999 Full Year Population Characteristics File) expands the scope of potential estimates. For example, to estimate the total number of office-based medical provider visits of persons with specific demographic characteristics (such as age, race, and sex), population characteristics from a person-level file need to be merged onto the office-based medical provider file. This procedure is illustrated below. The 1999 Appendix File provides examples of on how to merge MEPS other data files.

- 1) Create data set PERSX by sorting the 1999 Full Year Population Characteristics File, by the person identifier, DUPERSID. Keep only variables to be merged onto the office-based medical provider visit file and DUPERSID.
- 2) Create data set OBMP by sorting the office-based medical provider visit file by person identifier, DUPERSID.
- 3) Create final date set NEWOBMP by merging these two files by DUPERSID, keeping only records on the office-based medical provider visit file.

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

```
PROC SORT DATA=1999 Full Year Population Characteristics file
(KEEP=DUPERSID AGE SEX RACEX)
OUT=PERSX;
  BY DUPERSID;
RUN;

PROC SORT DATA=OBMP;
  BY DUPERSID;
RUN;

DATA NEWOBMP;
  MERGE OBMP (IN=A) PERSX(IN=B);
  BY DUPERSID;
  IF A;
RUN;
```

5.2 Linking the Office-Based Medical Provider Visit file to the MEPS 1999 Medical Conditions File and/or the MEPS 1999 Prescribed Medicines File

Due to survey design issues, there are limitations/caveats that data users/analyst must keep in mind when linking the different files. This limitations/caveats are listed below. For detailed linking examples, including SAS code, data users/analyst should refer to the 1999 Appendix File.

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

The RXLK file provides a link from the prescribed medicine records to the other event files. When using RXLK, data users/analysts should keep in mind that one office-based medical visit can link to more than one prescribed medicine record. Conversely, a prescribed medicine event may link to more than one office-based medical visits 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 Medical Conditions Link File)

The CLNK provides a link from MEPS event files to the 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 office-based medical provider visit. Users should also note that not all office-based medical provider visits 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, 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*.
- 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.
- Cox, B.G. and Cohen, S.B. (1985). Chapter 6: A Comparison of Household and Provider Reports of Medical Conditions. In *Methodological Issues for Health Care Surveys*. Marcel Dekker, New York.
- 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.
- Cox, B. and Iachan, R. (1987). A Comparison of Household and Provider Reports of Medical Conditions. Journal of the American Statistical Association 82(400):1013-18.
- Edwards, W.S., Winn, D.M., Kurlantzick V., et al. (1994). Evaluation of National Health Interview Survey Diagnostic Reporting. National Center for Health Statistics, Vital Health 2(120).
- Elixhauser A., Steiner C.A., Whittington C.A., and McCarthy E. Clinical Classifications for Health Policy Research: Hospital Inpatient Statistics, 1995. Healthcare Cost and Utilization Project, HCUP-3 Research Note. Rockville, MD: Agency for Health Care Policy and Research; 1998. AHCPR Pub. No. 98-0049.
- Health Care Financing Administration (1980). International Classification of Diseases, 9th Revision, Clinical Modification (ICD-CM). Vol. 1. (DHHS Pub. No. (PHS) 80-1260). DHHS: U.S. Public Health Services.
- Johnson, A.E. and Sanchez, M.E. (1993). Household and Medical Provider Reports on Medical Conditions: National Medical Expenditure Survey, 1987. Journal of Economic and Social Measurement. Vol. 19, 199-233.

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., 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.

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 (IN-SCOPE) 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

Survey Administration Variables

Variable	Description	Source
DUID	Dwelling unit ID (encrypted)	Assigned in sampling
PID	Person number (encrypted)	Assigned in sampling
DUPERSID	Sample person ID (DUID + PID)	Assigned in sampling
EVNTIDX	Event ID	Assigned in Sampling
EVENTRN	Event round number	CAPI derived
OBR2FLAG	Indicates whether or not a Panel 3 Round 2 event occurred in 1999	Constructed
FFEEIDX	Flat fee ID	CAPI derived
MPCELIG	MPC eligibility flag	CAPI derived
MPCDATA	MPC data flag	CAPI derived

Medical Provider Visits Variables

Variable	Description	Source
OBDATEYR	Event date – year	CAPI derived
OBDATEMM	Event date – month	CAPI derived
OBDATEDD	Event date – day	CAPI derived
SEETLKPV	Did P visit provider in person or telephone	MV01
REFERDBY	P referred for this visit another physician	MV02
SEEDOC	Did P talk to MD this visit/phone call	MV03
MEDPTYPE	Type of medical person P talked to on visit date	MV04
TIMESPNT	Time spent with doctor/medical person	MV05
DOCATLOC	Any MDs work at location where P saw provider	MV06
VSTCTGRY	Best category for care P received on visit date	MV07
VSTRELCN	This visit/phone call related to specific condition	MV08
PHYSTH	This visit did P have physical therapy	MV10
OCCUPTH	This visit did P have occupational therapy	MV10
SPEECHTH	This visit did P have speech therapy	MV10
CHEMOTH	This visit did P have chemotherapy	MV10
RADIATTH	This visit did P have radiation therapy	MV10
KIDNEYD	This visit did P have kidney dialysis	MV10
IVTHER	This visit did P have IV therapy	MV10
DRUGTRT	This visit did P have treatment for drug or alcohol	MV10
RCVSHOT	This visit did P receive an allergy shot	MV10
PSYCHOTH	Did P have psychotherapy/counseling	MV10
LABTEST	This visit did P have lab tests	MV11
SONOGRAM	This visit did P have sonogram or ultrasound	MV11
XRAYS	This visit did P have x-rays	MV11
MAMMOG	This visit did P have a mammogram	MV11
MRI	This visit did P have MRI	MV11
EKG	This visit did P have EKG or ECG	MV11

Variable	Description	Source
EEG	During this visit did P have EEG	MV11
RCVVAC	This visit did P receive a vaccination	MV11
ANESTH	During this visit did P receive anesthesia	MV11
OTHSVCE	This visit did P have other diagnostic tests/exams	MV11
SURGPROC	Was surgical procedure performed on P this visit	MV12
SURGNAME	Surgical procedure name in categories	MV13
MEDPRESC	Any medicines prescribed for P this visit	MV14
VAPLACE	VA Facility Flag	Constructed
OBICD1X	3-digit ICD-9 condition code	Edited
OBICD2X	3-digit ICD-9 condition code	Edited
OBICD3X	3-digit ICD-9 condition code	Edited
OBICD4X	3-digit ICD-9 condition code	Edited
OBPRO1X	2-digit ICD-9 procedure code	Edited
OBCCC1X	Modified Clinical Classification Code	Constructed/Edited
OBCCC2X	Modified Clinical Classification Code	Constructed/Edited
OBCCC3X	Modified Clinical Classification Code	Constructed/Edited
OBCCC4X	Modified Clinical Classification Code	Constructed/Edited

Flat Fee Variables

Variable	Description	Source
FFOBTYP	Edited Flat Bundle	FF01,FF02 (Edited)
FFBEF99	Total # visits in flat fee before 1999	FF05
FFTOT00	Number of visits in flat fee after 1999	FF10

Imputed Expenditure Variables

OBSF99X	Amount paid, family (imputed)	CP11 (Edited/Imputed)
OBNR99X	Amount paid, Medicare (imputed)	CP09 (Edited/Imputed)
OBMD99X	Amount paid, Medicaid (imputed)	CP07 (Edited/Imputed)
OBPV99X	Amount paid, Private Insurance (imputed)	CP07 (Edited/Imputed)
OBVA99X	Amount paid, Veterans (imputed)	CP07 (Edited/Imputed)
OBCH99X	Amount paid, CHAMP/CHAMPVA (imputed)	CP07 (Edited/Imputed)
OBOF99X	Amount paid, other federal (imputed)	CP07 (Edited/Imputed)
OBSL99X	Amount paid, state/local govt. (imputed)	CP07 (Edited/Imputed)
OBWC99X	Amount paid, Worker's Comp (imputed)	CP07 (Edited/Imputed)
OBOR99X	Amount paid, other private (imputed)	Constructed
OBOU99X	Amount paid, other public (imputed)	Constructed
OBOT99X	Amount paid, other insurance (imputed)	CP07 (Edited/Imputed)
OBXP99X	Sum of payments OBSF99X – OBOT99X	Constructed
OBTC99X	Total charge (imputed)	CP09 (Edited/Imputed)
IMPFLAG	Imputation Status	Constructed

Weights

Variable	Description	Source
PERWT99F	Final person level weight, 1999	Constructed
VARPSU99	Variance estimation PSU 1999	Constructed
VARSTR99	Variance estimation stratum	Constructed