MEPS HC-051G:
2000 Office-Based Medical Provider Visits

March 2003

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 micro-data contained in these files. Nevertheless, under sections 308 (d) and 903 (c) of the Public Health Service Act (42 U.S.C. 242m and 42 U.S.C. 299 a-1), data collected by the Agency for Healthcare Research and Quality (AHRQ) and/or the National Center for Health Statistics (NCHS) may not be used for any purpose other than for the purpose for which they were supplied; any effort to determine the identity of any reported cases is prohibited by law.

Therefore in accordance with the above referenced Federal Statute, it is understood that:

1. No one is to use the data in this data set in any way except for statistical reporting and analysis; and

2. If the identity of any person or establishment should be discovered inadvertently, then (a) no use will be made of this knowledge, (b) the Director, Office of Management, AHRQ will be advised of this incident, (c) the information that would identify any individual or establishment will be safeguarded or destroyed, as requested by AHRQ, and (d) no one else will be informed of the discovered identity; and

3. No one will attempt to link this data set with individually identifiable records from any data sets other than the Medical Expenditure Panel Survey or the National Health Interview Survey.

By using these data you signify your agreement to comply with the above stated statutorily based requirements with the knowledge that deliberately making a false statement in any matter within the jurisdiction of any department or agency of the Federal Government violates Title 18 part 1 Chapter 47 Section 1001 and is punishable by a fine of up to $10,000 or up to 5 years in prison.

The Agency for Healthcare Research and Quality requests that users cite AHRQ and the Medical Expenditure Panel Survey as the data source in any publications or research based upon these data.

B. BACKGROUND

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

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

MEPS is the third in a series of national probability surveys conducted by AHRQ on the financing and use of medical care in the United States. The National Medical Care Expenditure Survey (NMCES, also known as NMES-1) was conducted in 1977 and the National Medical Expenditure Survey (NMES-2) in 1987. Since 1996, MEPS continues this series with design enhancements and efficiencies that provide a more current data resource to capture the changing dynamics of the health care delivery and insurance 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 advance these goals, MEPS includes linkage with the National Health Interview Survey (NHIS) - a survey conducted by NCHS from which the sample for the MEPS HC is drawn - and enhanced longitudinal data collection for core survey components. The MEPS HC augments NHIS by selecting a sample of NHIS respondents, collecting additional data on their health care expenditures, and linking these data with additional information collected from the respondents' medical providers, employers, and insurance providers.

1.0 Household Component

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

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

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

2.0 Medical Provider Component

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

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

3.0 Insurance Component

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

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

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

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

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

4.0 Survey Management

MEPS data are collected under the authority of the Public Health Service Act. They are edited and published in accordance with the confidentiality provisions of this act and the Privacy Act. NCHS provides consultation and technical assistance.

As soon as data collection and editing are completed, the MEPS survey data are released to the public in staged releases of summary reports, microdata files and compendiums of tables. Data are released through MEPSnet, an online interactive tool developed to give users the ability to statistically analyze MEPS data in real time. Summary reports and compendiums of tables are released as printed documents and electronic files. Microdata files are released on electronic files.

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

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

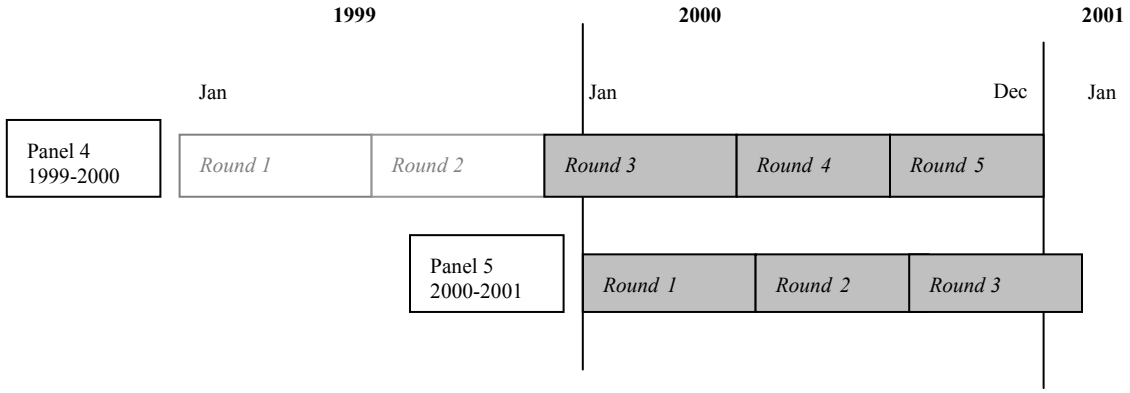
Be sure to specify the AHRQ number of the document you are requesting.

Additional information on MEPS is available from the MEPS project manager or the MEPS public use data manager at the Center for 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 2000 Medical Expenditure Panel Survey (MEPS) Household (HC) and Medical Provider Components (MPC). Released as an ASCII data file and a SAS transport file, the 2000 Office-Based Medical Provider Visits 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 2000. As illustrated below, this file consists of MEPS survey data from the 2000 portion of Round 3 and Rounds 4 and 5 for Panel 4, as well as Rounds 1, 2 and the 2000 portion of Round 3 for Panel 5 (i.e., the rounds for the MEPS panels covering calendar year 2000).



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. Utilization counts of office-based provider visits are based entirely on household reports. 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 2000 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 also be used to construct summary variables of expenditures, sources of payment, and related aspects of office-based provider visits for calendar year 2000. Aggregate annual person-level information on the use of office-based providers and other health services use is provided on the MEPS 2000 Full Year Consolidated Data File, where each record represents a MEPS sampled person.

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

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

For more information on MEPS HC survey design, see S. Cohen, 1997; J. Cohen, 1997; and S. Cohen, 1996. 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 2000 Office-Based Medical Provider public use data set consists of one event-level data file. The file contains characteristics associated with the OB event and imputed expenditure data. For users wanting to impute expenditures, pre-imputed data are available through the CCFS data center. Please visit the CCFS Data Denter website for details: <http://www.meps.ahrq.gov/>. The data user/analyst is forewarned that the imputation of expenditures will necessitate a sizable commitment of resources: financial, staff, and time.

The Office-Based Provider public use data set contains 102,530 office-based provider event records; of these records, 99,939 are associated with persons having a positive person-level weight (PERWT00F). 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 2000. Office-based provider visits known to have occurred after December 31, 2000 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 2000 portion of Round 3, and Rounds 4 and 5 for Panel 4, as well as Rounds

1, 2, and the 2000 portion of Round 3 for Panel 5 of the MEPS Household Component. 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 2000 eligibility (i.e., persons with a positive 2000 full-year person-level weight (PERWT00F > 0)), or

b) Be an eligible member of a family all of whose key in-scope members have a positive person-level weight (PERWT00F > 0). (Such a family consists of all persons with the same value for FAMIDYR.) That is, the person must have a positive full-year family-level weight (FAMWT00F >0). Note that FAMIDYR and FAMWT00F are variables on the 2000 Population Characteristics file.

Persons with no office-based medical provider visit events for 2000 are not included on this file but are represented on the 2000 MEPS person-level files. A codebook for the data file is provided in files H51GCB.PDF and H51GCB.ASP.

Each office-based medical provider visit 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 this file can be merged with the MEPS 2000 Full Year Population Characteristics file using the unique person identifier, DUPERSID, to append person-level information, such as demographic or health insurance characteristics, to each record. The office-based medical provider visit events can also be linked to the MEPS 2000 Medical Conditions File and MEPS 2000 Prescribed Medicines File. Please see section 5.0 for details on how to merge MEPS data files.

Panel 4 cases (PANEL00 = 4 on the MEPS 2000 Full Year Population Characteristics File) can also be linked back to the 1999 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 4 data.

2.1 Codebook Structure

For each variable on the office-based provider file, both weighted and unweighted frequencies are provided in the codebook (files H51GCB.PDF and H51GCB.ASP). The codebook and data file sequence list variables in the following order:

Unique person identifiers
Unique office-based medical provider visit event identifier
Office-based medical provider visit characteristic variables
ICD-9 condition and procedure codes
Clinical Classification Software (CCS) 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, values of -1, -7, -8, and -9 for non-expenditure variables have not been edited on this file. The values of -1 and -9 can be edited by the data users/analysts by following the skip patterns in the HC survey questionnaire (located on the MEPS web site: <http://www.meps.ahrq.gov/>).

2.3 Codebook Format

The office-based medical provider visits 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 Source and Naming Conventions

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 itself, derived from the MPC data collection instrument, derived from CAPI, or assigned in sampling. The source of each variable is identified in Section D "Variable - Source Crosswalk" in one of four ways:

1) Variables derived from CAPI or assigned in sampling are indicated as "CAPI derived" or "Assigned in sampling," respectively;

2) Variables which come from one or more specific questions have those questionnaire sections and question numbers indicated in the "Source" column; questionnaire sections are identified as:

    - MV - Office-Based Medical Provider Visits section
    - FF - Flat Fee section
    - CP - Charge Payment section;

3) Variables constructed from multiple questions using complex algorithms are labeled "Constructed" in the "Source" column; and

4) Variables that have been edited or imputed are so indicated.

2.4.2 Expenditure and Source of Payment Variables

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

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

The first two characters indicate the type of event:

IP - inpatient stay	OB - office-based visit
ER - emergency room visit	OP - outpatient visit
HH - home health 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
MR - Medicare	SL - State/local government
MD - Medicaid	WC – Workers' Compensation
PV - private insurance	OT - other insurance
VA - Veterans	OR - other private
TR - TRICARE	OU - other public
	XP - sum of payments

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

The fifth and sixth characters indicate the year (00). The last character indicates whether the variable is edited/imputed (X).

For example, OBSF00X is the edited/imputed amount paid by self or family for an office-based medical provider visit expenditure incurred in 2000.

2.5 File Contents

2.5.1 Survey Administration Variables

2.5.1.1 Person Identifiers (DUID, PID, DUPERSID)

The dwelling unit ID (DUID) is a five-digit random number assigned after the case was sampled for MEPS. The three-digit person number (PID) uniquely identifies each person within the dwelling unit. The eight-character variable DUPERSID uniquely identifies each person represented on the file and is the combination of the variables DUID and PID. For detailed information on dwelling units and families, please refer to the documentation for the 2000 Full Year Population Characteristics.

2.5.1.2 Record Identifiers (EVNTIDX, FFEEIDX)

EVNTIDX uniquely identifies each office-based medical provider visit event (i.e., each record on the office-based medical provider visits file) and is the variable required for linking office-based medical provider visit events to data files containing details on conditions and/or prescribed medicines (MEPS 2000 Medical Condition file and MEPS 2000 Prescribed Medicine file, respectively). For details on linking see Section 5.0 or the MEPS 2000 Appendix File, HC-051I.

FFEEIDX is a constructed variable that 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. A "mixed" flat fee group could contain both outpatient and office-based visits. Only outpatient and office-based events are allowed in a mixed bundle. Please note that FFEEIDX should be used to link up the outpatient and office-based events in order to determine the full set of events that are part of a flat fee group.

2.5.1.3 Record Indicator (EVENTRN)

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

2.5.2 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 Office-Based Medical Provider Visit Variables

The file contains variables describing office-based medical provider visit events reported by respondents in the Medical Provider Visits section of the MEPS HC survey questionnaire.

2.5.3.1 Date of Visit (OBDATEYR - OBDATEDD)

There are three variables that, together, indicate the day, month and year an office-based provider visit occurred (OBDATEDD, OBDATEMM, OBDATEYR, respectively). These variables have not been edited or imputed.

2.5.3.2 Visit Details (SEETLKPV-VSTRELCN)

The questionnaire determines if during the office-based medical provider visit 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.

2.5.3.3 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 “inapplicable,” “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).

2.5.3.4 VA Facility (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 "No".

2.5.4 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 and CCS codes (OBICD1X-OBICD4X, OBCCC1X-OBCCC4X) and one procedure code (OBPRO1X) 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 and procedures reported by the Household Component respondent were recorded by the interviewer as verbatim text, which were then coded to fully-specified 2000 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 3 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 how conditions and procedures were coded, please refer to the documentation on the MEPS 2000 Medical Conditions File. For frequencies of conditions by event type, please see the MEPS 2000 Appendix File, HC-051I.

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 three-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. Details on this procedure can be found in the 2000 MEPS Medical Conditions File.

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 order of input into the database 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 visits file should note that the order of conditions on this file is not identical to that on the 2000 Medical Conditions file.

2.5.5 Flat Fee Variables (FFEEIDX, FFOBTYPE, FFBEF00, FFTOT01)

2.5.5.1 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, and the associated 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 include flat fee groups where at least one of the health care events, as reported by the HC respondent, occurred during 2000. By definition, a flat fee group can span multiple years and/or event types (only outpatient department visits and physician office visits). Furthermore, a single person can have multiple flat fee groups.

2.5.5.2 Flat Fee Variable Descriptions

2.5.5.2.1 Flat Fee ID (FFEEIDX)

As noted earlier in Section 2.5.1.2 "Record Identifiers," the variable FFEEIDX uniquely identifies all events that are part of the same flat fee group for a person. On any 2000 MEPS event file, every event that is part of a specific flat fee group will have the same value for FFEEIDX. Note that prescribed medicine and home health events are never included in a flat fee group and FFEEIDX is not a variable on those event files.

2.5.5.2.2 Flat Fee Type (FFOBTYPE)

FFOBTYPE indicates whether the 2000 office-based medical provider visit event is the "stem" or "leaf" of a flat fee group. A stem (records with FFOBTYPE = 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 FFOBTYPE = 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. For the office-based visits that are not part of a flat fee payment situation, the FFOBTYPE is set to -1, "INAPPLICABLE."

2.5.5.2.3 Counts of Flat Fee Events that Cross Years (FFBEF00, FFTOT01)

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 the office-based medical provider visit occurred in 2000 as a part of a group of events, and some of the events occurred before 2000, counts of the known events are provided on the office-based medical provider visit event file record. Variables indicating events that occurred before or after 2000 as follows. These variables are:

> FFBEF00 - total number of pre-2000 events in the same flat fee group as the 2000 office-based medical provider visit. This count would not include the 2000 office-based medical visit(s).
>
> FFTOT01 - indicates the number of 2001 office-based events expected to be in the same flat fee group as the office-based medical provider visit event(s) that occurred in 2000.

2.5.5.3 Caveats of Flat Fee Groups

Data users/analysts should note that flat fee payment situations are common on the office-based medical provider visits file. There are 2,329 office-based medical provider visit 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 those in 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 2000, but the remaining visits that were part of this flat fee group occurred in 2001. In this case, the 2000 flat fee group represented on this file would consist of one event (the stem). The 2001 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 1999 but subsequent visits occurred during 2000. In this case, the initial visit would not be represented on the file. This 2000 flat fee group would then consist only of one or more leaf records and no stem. Another reason for which a flat fee group would not have a stem and at least one leaf record is that the stem or leaves could have been reported as different event types. Outpatient and Office-based medical provider visits are the only

two event types allowed in a single flat fee group. The stem may have been reported as an outpatient department visit and the leaves may have been reported as office-based medical provider visits.

2.5.6 Expenditure Data

2.5.6.1 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 is it directly comparable to the resource costs of those services or 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 the CCFS data center. For more information, see the Data Center section of the MEPS web site <http://www.meps.ahrq.gov>.

2.5.6.2 Data Editing and Imputation Methodologies of Expenditure Variables

The expenditure data included on this file were derived from both the MEPS household (HC) and medical provider (MPC) components. 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.

2.5.6.2.1 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.

2.5.6.2.2 General Hot-Deck Imputation

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

2.5.6.2.3 Office-Based Provider Visit Data Editing and Imputation

Facility expenditures for office-based provider visits were developed in a sequence of logical edits and imputations. "Household" edits were applied to sources and amounts of payment for all events reported by HC respondents. "MPC" edits were applied to provider-reported sources and amounts of payment for records matched to household-reported events. Both sets of edits were used to correct obvious errors (as described above) in the reporting of expenditures. After the data from each source were edited, a decision was made as to whether household- or MPC-reported information would be used in the final editing and hot-deck imputations for missing expenditures. The general rule was that MPC data would be used for events where a household-reported event corresponded to an MPC-reported event (i.e., a matched event), since providers usually have more complete and accurate data on sources and amounts of payment than households.

Separate imputations were performed for flat fee and simple events.

Logical edits also 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 the extent of its missing charge and expenditure data. For example, an event with a known total charge but no expenditure information was assigned to one category, while an event with a known total charge and partial expenditure information was assigned to a different category. Similarly, events without a known total charge and no or partial expenditure information were assigned to various recipient categories.

The logical edits produced eight recipient categories in which all events had a common extent of missing data. Separate hot-deck imputations were performed on events in each recipient category, and the donor pool was restricted to events with complete expenditures from the MPC. The donor pool restriction was used even though some unmatched events had complete household-reported expenditures. These events were not allowed to donate information to other events because the MPC data were considered to be more reliable.

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 both implicitly included in paid events, and explicitly included in events that should have been treated as free from provider.

2.5.6.3 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.

2.5.6.4 Imputation Flag (IMPFLAG)

Only one imputation flag was created for 2000 event files. This variable, IMPFLAG, is a six-category variable that indicates if the event contains complete Household Component (HC) or Medical Provider Component (MPC) data, was fully or partially imputed, or was

imputed in the capitated imputation process (for OP and MV events only). The following list identifies how the imputation flag is coded; the categories are mutually exclusive.

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

IMPFLAG=1 complete HC data

IMPFLAG=2 complete MPC data

IMPFLAG=3 fully imputed

IMPFLAG=4 partially imputed

IMPFLAG=5 complete MPC data through capitation imputation

2.5.6.5 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 2000, all of the events that occurred in 2000 will have zero payments. Conversely, if the first event in the flat fee group occurred at the end of 2000, the total expenditure for the entire flat fee group will be on that event, regardless of the number of events it covered after 2000. See Section 2.5.3 for details on the flat fee variables.

2.5.6.6 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.

2.5.6.7 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.

2.5.6.8 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 (self) or family
2. Medicare
3. Medicaid
4. Private Insurance
5. Veterans Administration, excluding TRICARE
6. TRICARE, formerly CHAMPUS/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. Workers' Compensation
10. Other Unclassified Sources - includes sources such as automobile, homeowner's, liability, and other miscellaneous or unknown sources.

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

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

Though relatively small in magnitude, data users/analysts should exercise caution when interpreting the expenditures associated with these two additional sources of payment. While these payments stem from apparent inconsistent responses to health insurance and source of payment questions in the survey, some of these inconsistencies may have logical explanations. For example, private insurance coverage in MEPS is defined as having a major medical plan covering hospital and physician services. If a MEPS sampled person did not have such coverage but had a single service type insurance plan (e.g., dental insurance) that paid for a particular episode of care, those payments may be classified as "other private". Some of the "other public" payments may stem from confusion between Medicaid and other state and local programs or may be persons who were not enrolled in Medicaid, but were presumed eligible by a provider who ultimately received payments from the program.

2.5.6.9 Office-Based Expenditure Variables (OBSF00X - OBTC00X)

OBSF00X - OBOT00X are the 12 sources of payment. OBTC00X is the total charge, and OBXP00X is the sum of the 12 sources of payment for the office-based provider visit expenditures. The 12 source of payment are: self/family (OBSF00X), Medicare (OBMR00X), Medicaid (OBMD00X), private insurance (OBPV00X), Veterans Administration (OBVA00X), TRICARE (OBTR00X), other Federal sources (OBOF00X), State and Local (non-federal) government sources (OBSL00X), Worker's Compensation (OBWC00X), other private insurance (OBOR00X), other public insurance (OBOU00X), and other insurance (OBOT00X).

2.5.7 Rounding

Expenditure variables have been rounded to the nearest penny. Person-level expenditure information released on the MEPS 2000 Person-Level Expenditure File will be rounded to the nearest dollar. It should be noted that using the MEPS 2000 event files to create person-level totals will yield slightly different totals than that those found on the person-level expenditure file. These differences are due to rounding only. Moreover, in some instances, the number of persons having expenditures on the event files for a particular source of payment may differ from the number of persons with expenditures on the person-level expenditure file for that source of payment. This difference is also an artifact of rounding only. Please see the MEPS 2000 Appendix File, HC-051I, for details on such rounding differences.

3.0 Sample Weight (PERWT00F)

3.1 Overview

There is a single full year person-level weight (PERWT00F) 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 2000. 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 Weight Construction

The person-level weight PERWT00F was developed in several stages. Person level weights for Panels 4 and 5 were created separately. The weighting process for each panel included an adjustment for nonresponse over time and poststratification. Poststratification was achieved by 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. A 2000 composite weight was then formed by multiplying each panel weight by .5 and then poststratifying the resulting weight to the same set of CPS-based control totals. When poverty status information derived from income variables became available, a final poststratification was done on the resulting 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 4 Weight

The person-level weight for MEPS Panel 4 was developed using the 1999 full year weight for an individual as a "base" weight for survey participants present in 1999. For key, in-scope respondents who joined an RU some time in 2000 after being out-of-scope in 1999, the 1999 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 2000. These control figures were derived by scaling back the population totals obtained from the March 2000 CPS to reflect the December 2000 CPS estimated population distribution across age and sex categories as of December 2000. 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, 2000 is 275,158,755. Key, responding persons not in-scope on December 31, 2000 but in-scope earlier in the year retained, as their final Panel 4 weight, the weight after the nonresponse adjustment.

3.2.2 MEPS Panel 5 Weight

The person-level weight for MEPS Panel 5 was developed using the MEPS Round 1 person-level weight as a "base" weight. For key, in-scope respondents who joined an RU after Round 1, the Round 1 family weight served as a "base" weight. The weighting process included an adjustment for nonresponse over Round 2 and the 2000 portion of Round 3 as well as poststratification to the same population control figures for December 2000 used for the MEPS Panel 4 weights. The same five variables employed for Panel 4 poststratification (census region, MSA status, race/ethnicity, sex, and age) were used for Panel 5 poststratification. Similarly, for Panel 5, key, responding persons not in-scope on December 31, 2000 but in-scope earlier in the year retained, as their final Panel 5 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 2000 CPS data base.

3.2.3 The Final Weight for 2000

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, 2000 is 275,158,755 (PERWT00F>0 and INSC1231=1). The weights of some persons out-of-scope on December 31, 2000 were also poststratified. Specifically, the weights of persons out-of-scope on December 31, 2000 who were inscope some time during the year and also entered a nursing home during the year were poststratified to a corresponding control total obtained from the 1996 MEPS Nursing Home Component. The weights of persons who died while inscope during 2000 were poststratified to corresponding estimates derived using data obtained from the Medicare

Current Beneficiary Survey (MCBS) and Vital Statistics information provided by the National Center for Health Statistics (NCHS). Separate control totals were developed for the "65 and older" and "under 65" civilian, noninstitutionalized populations.

3.2.4 Coverage

The target population for MEPS in this file is the 2000 U.S. civilian noninstitutionalized population. However, the MEPS sampled households are a subsample of the NHIS households interviewed in 1998 (Panel 4) and 1999 (Panel 5). New households created after the NHIS interviews for the respective Panels and consisting exclusively of persons who entered the target population after 1998 (Panel 4) or after 1999 (Panel 5) are not covered by MEPS. Neither are previously out of scope persons who join an existing household but are unrelated to the current household residents. Persons not covered by a given MEPS panel thus include some members of the following groups: immigrants; persons leaving the military; U.S. citizens returning from residence in another country; and persons leaving institutions. The set of uncovered persons constitutes only a small proportion of the MEPS target population.

4.0 Strategies for Estimation

This file is constructed for efficient estimation of utilization, expenditures, and sources of payment for office-based medical provider visits and to allow for estimates of number of persons with office-based medical provider visits in 2000.

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, Expenditures, and Sources of Payment

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

In order to produce national estimates related to office-based medical provider visits utilization, expenditures, and sources of payment, the value in each record contributing to the estimates must be multiplied by the weight (PERWT00F) contained on that record.

Example 1

For example, the total number of office-based medical provider visits for the civilian noninstitutionalized population of the U.S. in 2000 is estimated as the sum of the weight (PERWT00F) across all office-based medical provider visit records. That is,

$$\Sigma W_j = 1{,}236{,}897{,}482 \qquad (1)$$

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 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,

$$(\sum W_j X_j)/(\sum W_j) = \$19.58 \qquad (2)$$

where

$\sum W_j = 1{,}145{,}624{,}853$ and $X_j = OBSF00X_j$

for all records with $OBXP00X_j > 0$.

This gives $19.58as the estimated mean amount of out-of-pocket payment of expenditures associated with office-based medical provider visits and 1,145,624,853 as an estimate of the total number of office-based medical provider visits with expenditure. Both of these estimates are for the civilian noninstitutionalized population of the U.S. in 2000.

Example 3

Another example would be to estimate the mean 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 the proportion of total expenditures paid by private insurance at the provider visit level. That is

$$(\sum W_j Y_j)/(\sum W_j) = 0.4373 \qquad (3)$$

where

$\sum W_j = 1{,}145{,}624{,}853$ and $Y_j = OBPV00X_j / OBXP00X_j$

for all office-based medical provider visits with $OBXP00X_j > 0$.

This gives 0.4373 as the estimated mean proportion of total expenditures paid by private insurance for office-based medical provider visits with expenditures for the civilian noninstitutionalized population of the U.S. in 2000.

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

When calculating an estimate of the total number of persons with office-based medical provider visits, users can use a person-level file 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,

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

where

W_i is the sampling weight (PERWT00F) for person i

and

$X_i = 1$ if $SEETLKPV_j = 1$ 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 Visit Events

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,

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

where

W_i is the sampling weight (PERWT00F) for person i

and

$Z_i = \sum OBXP00X_j$ 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, the Full Year Consolidated 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 the civilian noninstitutionalized 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,

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

where

W_i is the sampling weight (PERWT00F) 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 (VARSTR00, VARPSU00)

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 2000 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 VARSTR00 and VARPSU00, 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 VARSTR00 and VARPSU00 as the variance estimation strata and PSUs (within these strata), respectively, and specifying a "with replacement" design in a computer software package (i.e., SUDAAN) will yield standard error estimates of $0.5677 and 0.0087 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 visits 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 2000 Person-Level File to the 2000 Office-Based Medical Provider Visits File

Merging characteristics of interest from person-level files (e.g., MEPS 2000 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 visits file. This procedure is illustrated below. The MEPS 2000 Appendix File, HC-051I provides additional detail on how to merge MEPS data files.

1) Create data set PERSX by sorting the 2000 Full Year Population Characteristics File, by the person identifier, DUPERSID. Keep only variables to be merged onto the office-based medical provider visits file and DUPERSID.

2) Create data set OBMP by sorting the office-based medical provider visits file by person identifier, DUPERSID.

3) Create final data set NEWOBMP by merging these two files by DUPERSID, keeping only records on the office-based medical provider visits file.

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

```
PROC SORT DATA=2000 Full Year Population Characteristics file
(KEEP=DUPERSID AGE31X AGE42X AGE53X 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 MEPS 2000 Office-Based Medical Provider Visits File to the MEPS 2000 Medical Conditions File and/or the MEPS 2000 Prescribed Medicines File

Due to survey design issues, there are limitations/caveats that data users/analysts must keep in mind when linking the different files. These limitations/caveats are listed below. For detailed linking examples, including SAS code, data users/analysts should refer to the MEPS 2000 Appendix File, HC-051I.

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

The RXLK file provides a link from MEPS event files to the 2000 Prescribed Medicine File. 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.

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

The CLNK provides a link from MEPS event files to the 2000 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 an 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. (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. 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., 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.

D. VARIABLE-SOURCE CROSSWALK

VARIABLE-SOURCE CROSSWALK

FOR MEPS HC-051G: 2000 OFFICE-BASED MEDICAL PROVIDER VISITS

Survey Administration Variables

Variable	Description	Source
DUID	Dwelling unit ID	Assigned in sampling
PID	Person number	Assigned in sampling
DUPERSID	Person ID (DUID + PID)	Assigned in sampling
EVNTIDX	Event ID	Assigned in Sampling
EVENTRN	Event round number	CAPI derived
FFEEIDX	Flat fee ID	CAPI derived
MPCELIG	MPC eligibility flag	Constructed
MPCDATA	MPC data flag	Constructed

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 Person visit provider in person or telephone	MV01
REFERDBY	This visit referred by another physician	MV02
SEEDOC	Did Person talk to MD this visit/phone call	MV03
MEDPTYPE	Type of medical person Person talked to on visit date	MV04
TIMESPNT	Time Person spent with doctor/medical person	MV05
DOCATLOC	Any MD work at location where Person saw provider	MV06
VSTCTGRY	Best category for care Person received on visit date	MV07
VSTRELCN	Was this visit/phone call related to spec hlth cond	MV08

Variable	Description	Source
PHYSTH	This visit did Person have physical therapy	MV10
OCCUPTH	This visit did Person have occupational therapy	MV10
SPEECHTH	This visit did Person have speech therapy	MV10
CHEMOTH	This visit did Person have chemotherapy	MV10
RADIATTH	This visit did Person have radiation therapy	MV10
KIDNEYD	This visit did Person have kidney dialysis	MV10
IVTHER	This visit did Person have IV therapy	MV10
DRUGTRT	This visit did Person have treatment for drug/alcohol	MV10
RCVSHOT	This visit did Person receive an allergy shot	MV10
PSYCHOTH	Did Person have psychotherapy/counseling	MV10
LABTEST	This visit did Person have lab tests	MV11
SONOGRAM	This visit did Person have sonogram or ultrasound	MV11
XRAYS	This visit did Person have x-rays	MV11
MAMMOG	This visit did Person have a mammogram	MV11
MRI	This visit did Person have an MRI/Catscan	MV11
EKG	This visit did Person have an EKG or ECG	MV11
EEG	This visit did Person have an EEG	MV11
RCVVAC	This visit did Person receive a vaccination	MV11
ANESTH	This visit did Person receive anesthesia	MV11
OTHSVCE	This visit did Person have other diagnostic test/exam	MV11
SURGPROC	Was surgical procedure performed on Person this visit	MV12
SURGNAME	Surgical procedure name in categories	MV13
MEDPRESC	Any medicines prescribed for Person 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

Variable	Description	Source
OBICD4X	3-digit ICD-9 condition code	Edited
OBPRO1X	2-digit ICD-9 procedure code	Edited
OBCCC1X	Modified Clinical Classification Code (CCS)	Constructed/Edited
OBCCC2X	Modified Clinical Classification Code (CCS)	Constructed/Edited
OBCCC3X	Modified Clinical Classification Code (CCS)	Constructed/Edited
OBCCC4X	Modified Clinical Classification Code (CCS)	Constructed/Edited

Flat Fee Variables

Variable	Description	Source
FFOBTYPE	Flat fee bundle	Constructed
FFBEF00	Total # of visits in FF before 2000	FF05
FFTOT01	Total # of visits in FF after 2000	FF10

Imputed Expenditure Variables

Variable	Description	Source
OBSF00X	Amount paid, self/family (imputed)	CP Section (Edited)
OBMR00X	Amount paid, Medicare (imputed)	CP Section (Edited)
OBMD00X	Amount paid, Medicaid (imputed)	CP Section (Edited)
OBPV00X	Amount paid, private insurance (imputed)	CP Section (Edited)
OBVA00X	Amount paid, Veterans Administration (imputed)	CP Section (Edited)
OBTR00X	Amount paid, TRICARE (imputed)	CP Section (Edited)
OBOF00X	Amount paid, other federal (imputed)	CP Section (Edited)
OBSL00X	Amount paid, state & local government (imputed)	CP Section (Edited)
OBWC00X	Amount paid, workers' compensation (imputed)	CP Section (Edited)
OBOR00X	Amount paid, other private insurance (imputed)	Constructed
OBOU00X	Amount paid, other public insurance (imputed)	Constructed
OBOT00X	Amount paid, other insurance (imputed)	CP Section (Edited)

Variable	Description	Source
OBXP00X	Sum of OBSF00X – OBOT00X (imputed)	Constructed
OBTC00X	Household reported total charge (imputed)	CP Section (Edited)
IMPFLAG	Imputation status	Constructed

Weights

Variable	Description	Source
PERWT00F	Final person level weight, 2000	Constructed
VARSTR00	Variance estimation stratum, 2000	Constructed
VARPSU00	Variance estimation PSU, 2000	Constructed