MEPS HC-085D:
2004 Hospital Inpatient Stays

September 2006

Agency for Healthcare Research and Quality
Center for Financing, Access and Cost Trends
540 Gaither Road
Rockville, MD 20850
(301) 427-1406

Table of Contents

A. Data Use Agreement

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

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

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

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

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

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

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

B. Background

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

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

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

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

1.0 Household Component

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

The HC uses an overlapping panel design in which data are collected through a preliminary contact followed by a series of five rounds of interviews over a 2 ½-year

period. Using computer-assisted personal interviewing (CAPI) technology, data on medical expenditures and use for two calendar years are collected from each household. This series of data collection rounds is launched each subsequent year on a new sample of households to provide overlapping panels of survey data and, when combined with other ongoing panels, will provide continuous and current estimates of health care expenditures.

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

2.0 Medical Provider Component

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

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

3.0 Insurance Component

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

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

- A list of employers or other insurance providers identified by MEPS HC respondents who report having private health insurance at the Round 1 interview.

- A Bureau of the Census list frame of private sector business establishments.

- The Census of Governments from Bureau of the Census.

To provide an integrated picture of health insurance, data collected from the first sampling frame (employers and insurance providers identified by MEPS HC respondents) are linked back to data provided by those respondents. Data from the two Census Bureau sampling frames are used to produce annual national and state estimates

of the supply and cost of private health insurance available to American workers and to evaluate policy issues pertaining to health insurance. National estimates of employer contributions to group insurance from the MEPS IC are used in the computation of Gross Domestic Product (GDP) by the Bureau of Economic Analysis.

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

4.0 Survey Management

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

As soon as data collection and editing are complete, 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 CD-ROM and/or as electronic files.

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

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

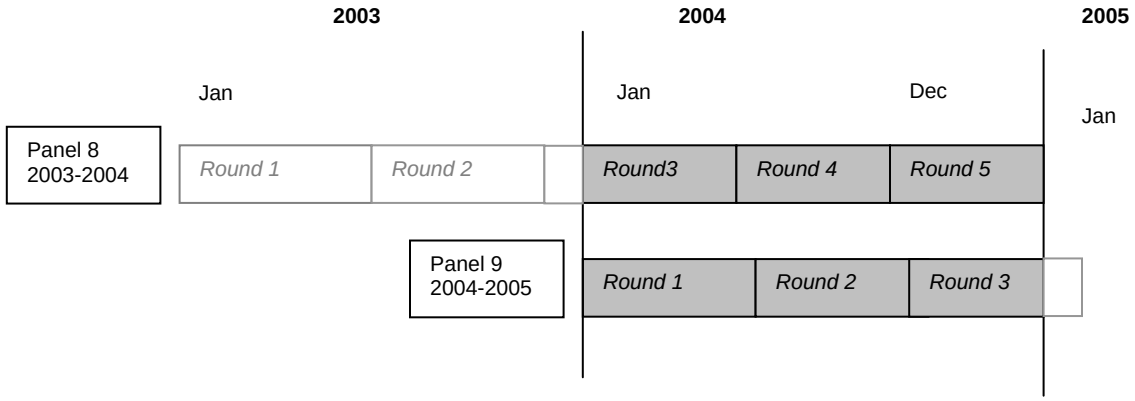
Be sure to specify the AHRQ number of the document you are requesting.

Additional information on MEPS is available from the MEPS project manager or the MEPS public use data manager at the Center for Financing, Access and Cost Trends, Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850 (301-427-1406).

C. Technical and Programming Information

1.0 General Information

This documentation describes one in a series of public use event files from the 2004 Medical Expenditure Panel Survey (MEPS) Household Component (HC) and Medical Provider Component (MPC). Released as an ASCII data file (with related SAS and SPSS programming statements) and SAS transport file, the 2004 Hospital Inpatient Stays (STAZ) public use file provides detailed information on hospital inpatient stays for a nationally representative sample of the civilian noninstitutionalized population of the United States. Data from the STAZ event file can be used to make estimates of hospital inpatient stay utilization and expenditures for calendar year 2004. The file contains 67 variables and has a logical record length of 371with an additional 2-byte carriage return/line feed at the end of each record. As illustrated below, this file consists of MEPS survey data from the 2004 portion of Round 3 and Rounds 4 and 5 for Panel 8, as well as Rounds 1, 2 and the 2004 portion of Round 3 for Panel 9 (i.e., the rounds for the MEPS panels covering calendar year 2004).



Hospital stay events reported in Panel 9 Round 3 and known to have begun after December 31, 2004 are not included on this file.

Each record on the inpatient hospital event file represents a unique hospital inpatient stay, that is, a hospital inpatient stay reported by the household respondent. In addition to expenditures related to the stay, each record contains household-reported medical conditions and procedures associated with the hospitalization and information on the length of stay.

Annual counts of hospital inpatient stay utilization are based entirely on household reports. Information from the MEPS MPC is used to supplement expenditure and payment data reported by the household and does not affect use estimates.

Data from this event file can be merged with other 2004 MEPS HC data files for purposes of appending person-level data such as demographic characteristics or health insurance coverage to each hospital inpatient stay record.

This file can also be used to construct summary variables of expenditures, sources of payment, and related aspects of hospital inpatient care. Aggregate annual person-level information on the use of hospital inpatient stays and other health services use is provided on the MEPS 2004 Full Year Consolidated Data File, where each record represents a MEPS sampled person.

This documentation offers an 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

Any variables not found on this file but released on previous years' files were excluded because they contained only missing data.

For more information on MEPS HC survey design see S. Cohen, 1997; J. Cohen, 1997; and S. Cohen, 1996. For information on the MEPS MPC design, see S. Cohen, 1999. Copies of the HC and the MPC survey instruments used to collect the information on the STAZ file are available in the *Survey Instruments* section on the MEPS web site at the following address: http://www.meps.ahrq.gov.

2.0 Data File Information

The 2004 Hospital Inpatient Stays public use data set consists of one event-level data file. The file contains characteristics associated with the STAZ event and imputed expenditure data.

The 2004 STAZ public use data set contains variable and frequency distributions for a total of 3,456 hospital inpatient stay records reported during the 2004 portion of Round 3 and Rounds 4 and 5 for Panel 8, as well as Rounds 1, 2, and the 2004 portion of Round 3 for Panel 9 of the MEPS Household Component. This file includes hospital inpatient stay records for all household survey respondents who resided in eligible responding households and reported at least one hospital inpatient stay. Hospital inpatient stay records known to have ended before January 1, 2004 or after December 31, 2004 are not included on this file. Some household respondents may have multiple hospital inpatient stays and, thus, will be represented in multiple records on this file. Other household

respondents may have reported no hospital inpatient stays and, thus, will have no records on this file. Of the 3,456 hospital inpatient stay records, 3,304 are associated with persons having a positive person-level weight (PERWT04F). The persons represented on this file had to meet the following three criteria:

1) The hospital stay had to have been reported by a household survey respondent as an inpatient hospital stay (regardless of a stay's length). Thus, the file contains some hospitalizations that were reported as not including an overnight stay.

2) The hospital stay had to have ended during 2004. Stays that began prior to 2004 but ended during 2004 are included on this data file; H85DCB.PDF provides the file codebook. Stays that began in 2004 but ended during 2005 are excluded from this data file and will be included in a subsequent 2005 IP data file. Persons with no hospital inpatient stay events for 2004 are not included on this event-level IP file but are represented on the person-level 2004 Full Year Population Characteristics file.

3) The persons represented on this file also had to meet either 3a or 3b:

a) Be classified as a key in-scope person who responded for his or her entire period of 2004 eligibility (i.e., persons with a positive 2004 full-year person-level sampling weight (PERWT04F > 0)), or

b) Be an eligible member of a family all of whose key in-scope members have a positive person-level weight (PERWT04F > 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 (FAMWT04F > 0). Note that FAMIDYR and FAMWT04F are variables on the 2004 Population Characteristics file.

One caveat that should be noted is that in the case of a newborn and the hospital inpatient stay associated with the newborn's birth, a separate hospital inpatient stay record exists on the file only if the newborn was discharged after the mother. Thus, hospital stays associated with a normal birth are generally represented on the file as a single record (i.e., the mother's hospital inpatient stay record, covering expenditure data for both the mother and baby). In situations where the newborn was discharged after the mother, the birth event will be represented as two records: one record for the mother and one record for the baby. For newborns re-admitted to the hospital during the reference year, each subsequent re-admission will have a separate record.

Each inpatient record includes the following: start and end dates of the hospital inpatient stay; number of nights in the hospital; reason entered the hospital; condition(s) associated with the hospital inpatient stay; medicines prescribed at discharge; flat fee information; imputed sources of payment; total payment and total charge for both the facility and physician portions of the hospital inpatient stay expenditure; a full-year person-level weight; variance strata; and variance PSU.

Data from this file can be merged with the MEPS 2004 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. Hospital inpatient stay events can also be linked to the MEPS 2004 Medical Conditions File and the MEPS 2004 Prescribed Medicines File. Please see Section 5.0 or the MEPS 2004 Appendix File, HC-085I, for details on how to merge MEPS data files.

2.1 Using MEPS Data for Trend and Longitudinal Analysis

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

The records on this file can be linked to all other 2004 MEPS-HC public use data sets by the sample person identifier (DUPERSID).

2.2 Codebook Structure

For each variable on the Inpatient Events file, both weighted and unweighted frequencies are provided in the codebook file (H85DCB.PDF). The codebook and data file sequence list variables in the following order:

Unique person identifiers
Unique hospital inpatient stay identifiers
Hospital inpatient stay characteristics variables
ICD-9-CM condition and procedure codes
Clinical Classification Software (CCS) codes
Imputed expenditure variables
Weight and variance estimation variables

2.3 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, the 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 data users/analysts by following the skip patterns in the HC survey questionnaire (located on the MEPS web site: http://www.meps.ahrq.gov).

2.4 Codebook Format

The STAZ 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

2.5 Variable Source and Naming Conventions

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

2.5.1 General

Variables on this file were derived from the HC 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:

- HS - Hospital Stays 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 which have been edited or imputed are so indicated.

2.5.2 Expenditure and Source of Payment Variables

The names of the expenditure and source of payment variables follow a standard convention, are eight 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 12 sources of payment variables 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

For expenditure variables on the IP file, the third character indicates whether the expenditure is associated with the facility (F) or the physician (D).

In the case of the source of payment variables, the fourth and fifth 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 sixth and seventh characters indicate the year (04). The eighth character, "X", indicates whether the variable is edited/imputed.

For example, IPFSF04X is the edited/imputed amount paid by self or family for the facility portion of the hospital inpatient stay expenditure incurred in 2004.

2.6 File Contents

2.6.1 Survey Administration Variables

2.6.1.1 Person Identifiers (DUID, PID, DUPERSID)

The dwelling unit ID (DUID) is a five-digit random number assigned after the case was sampled for MEPS. The three-digit person number (PID) uniquely identifies each person within the dwelling unit. The eight-character variable DUPERSID uniquely identifies each person represented on the file and is the combination of the variables DUID and PID. For detailed information on dwelling units and families, please refer to the documentation for the 2004 Full Year Population Characteristics File.

2.6.1.2 Record Identifiers (EVNTIDX, ERHEVIDX, FFEEIDX)

EVNTIDX uniquely identifies each hospital inpatient stay/event (i.e., each record on the STAZ file) and is the variable required to link hospital inpatient stay events to data files

containing details on conditions and/or prescribed medicines (MEPS 2004 Medical Conditions File and MEPS 2004 Prescribed Medicines File, respectively). For details on linking, see Section 5.0 or the MEPS 2004 Appendix File, HC-085I.

ERHEVIDX is a constructed variable identifying a STAZ record that includes the facility expenditures for the preceding emergency room visit. This variable is derived from provider-reported information on linked emergency room and inpatient stay events that matched to corresponding events reported by the household. The variable ERHEVIDX contains the EVNTIDX of the linked event. On the 2004 STAZ file, there are 562 hospital stays linked to a preceding emergency room visit, that is, there are records with a valid ERHEVIDX value. ERHEVIDX has not been reconciled with the unedited variable EMERROOM. Please note that, the physician expenditures associated with the emergency room visit remain on the emergency room file.

FFEEIDX is a constructed variable which uniquely identifies a flat fee group, that is, all events that were a part of a flat fee payment. For example, dialysis treatments are typically covered in a flat fee arrangement where all visits are covered under one flat fee dollar amount. These events would have the same value for FFEEIDX.

2.6.1.3 Round Indicator (EVENTRN)

EVENTRN indicates the round in which the hospital inpatient stay was first reported. Please note that Rounds 3, 4, and 5 are associated with MEPS survey data collected from Panel 8. Likewise, Rounds 1, 2, and 3 are associated with data collected from Panel 9.

2.6.2 MPC Data Indicator (MPCDATA)

MPCDATA is a constructed variable which indicates whether or not MPC data were collected for the hospital inpatient stay. While all hospital inpatient events are sampled into the Medical Provider Component, not all hospital inpatient stay records have MPC data associated with them. This is dependent upon the cooperation of the household respondent to provide permission forms to contact the hospital as well as the cooperation of the hospital to participate in the survey.

2.6.3 Hospital Inpatient Stay Event Variables

This file contains variables describing hospital inpatient stays/events reported by household respondents in the Hospital Stays section of the MEPS HC questionnaire. The questionnaire contains specific probes for determining details about the hospital inpatient stay.

2.6.3.1 Start and End Dates of Event (IPBEGDD-IPENDYR)

This file contains variables describing hospital inpatient stays reported by household respondents in the Hospital Section of the MEPS HC questionnaire. There are three variables which indicate the day, month, and year a hospital stay began (IPBEGDD, IPBEGMM, IPBEGYR, respectively). Similarly, there are three variables which indicate the day, month, and year a hospital stay ended (IPENDDD, IPENDMM, IPENDYR, respectively). These variables have not been edited.

2.6.3.2 Length of Stay (NUMNIGHX, NUMNIGHT)

NUMNIGHX denotes the length of a hospital inpatient stay. For stays beginning in 2003 and ending in 2004, this variable would include the nights associated with the entire visit. It was edited using the above mentioned begin and end dates of the hospital inpatient stay (Section 2.6.3.1). If the dates were unknown, then NUMNIGHX used the number from the unedited variable NUMNIGHT (number of nights in the hospital). If both the dates and NUMNIGHT were missing data, then NUMNIGHX was imputed. Users should note that NUMNIGHT was only asked for events with missing date information. Hence, it contains large amounts of missing data and cannot be used alone but rather in conjunction with date information.

2.6.3.3 Preceding ER Visits (EMERROOM)

The variable EMERROOM was derived directly from the Hospital Inpatient Stays section of the HC survey instrument and is unedited. EMERROOM describes whether or not the hospital inpatient stay began with an emergency room visit. Data users/analysts should be aware that no attempt was made to reconcile EMERROOM with information from the Emergency Room Visit File. Furthermore, no attempt has been made to reconcile the unedited EMERROOM variable with the edited ERHEVIDX variable (see 2.6.1.2).

2.6.3.4 Other Visit Detail (SPECCOND - VAPLACE)

Also provided are the following unedited variables: hospital inpatient stays related to a medical condition (SPECCOND); the reason the person entered the hospital (RSNINHOS); any operation or surgery performed while the respondent was in the hospital (ANYOPER).

With respect to RSNINHOS, please note that while there were 472 cases where RSNINHOS = 4 (reason entered hospital – to give birth to a baby), this does not mean that there were actually 472 *new births*. In fact, it may have been reported that the mother went to the hospital for delivery (hence, the interviewer would have assigned the event RSNINHOS = 4), but the mother could have had, for example, false labor pains or

a stillbirth. Thus, this unedited self-reported variable may be inconsistent with reported number of births (see the 2004 Full Year Population Characteristics File, section 2.6.2 "Navigating the MEPS Data with Information on Person Disposition Status"). In addition, RSNINHOS has not been reconciled with the ICD-9-CM condition codes, the procedure codes, or the CCC codes that are on the file; thus, there is a general inconsistency between reported reason in hospital and reported conditions.

VAPLACE is a constructed variable that indicates whether the service was provided 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".

2.6.3.5 Condition and Procedure Codes (IPICD1X-IPICD4X, IPPRO1X, IPPRO2X), and Clinical Classification Codes (IPCCC1X-IPCCC4X)

Information on household reported medical conditions and procedures associated with each hospital inpatient stay event is provided on this file. There are up to four condition and CCC codes (IPICD1X-IPICD4X, IPCCC1X-IPCCC4X) and up to two procedure codes (IPPRO1X and IPPRO2X) listed for each hospital inpatient stay event. In order to obtain complete condition information associated with an event, the data user/analyst must link to the MEPS 2004 Medical Conditions File. Details on how to link the 2004 STAZ file to the MEPS 2004 Medical Conditions File are provided in Section 5.2 and the MEPS 2004 Appendix File, HC-085I. The data user/analyst should note that because of 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 2004 ICD-9-CM codes, including medical condition and V codes (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 (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 2004 Medical Conditions File. For frequencies of conditions by event type, please see the MEPS 2004 Appendix File, HC-085I.

The ICD-9-CM condition codes were aggregated into clinically meaningful categories. These categories, included on the file as IPCCC1X-IPCCC4X, 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 263 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 three-digit categories. Similarly, the procedure codes have been collapsed from fully-specified codes to two-digit code categories. Because of this collapsing, it is possible for there to be duplicate ICD-9-CM condition or procedure codes linked to a single medical event when different fully-specified codes are collapsed into the same code. For more information on ICD-9-CM codes, see the HC-087 documentation.

The condition (and clinical classification codes) and procedure codes linked to each hospital inpatient stay 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 users/analysts who use the MEPS 2004 Medical Conditions File in conjunction with this hospital inpatient stay event file should note that the order of conditions on this file is not identical to that on the Medical Conditions file.

The user should also note that because of the design of the HC survey instrument, most hospital stays that are reported as being for a delivery (RSNINHOS=4) link to condition codes that are for pregnancy rather than a delivery. In addition, RSNINHOS has not been reconciled with the ICD-9-CM condition codes, the procedure codes, or the CCC codes that are on the file.

2.6.3.6 Discharge Detail (DSCHPMED)

DSCHPMED is derived directly from the Hospital Stays Section of the HC survey instrument. DSCHPMED indicates whether or not any medicines were prescribed at discharge.

2.6.4 Flat Fee Variables (FFEEIDX, FFIPTYPE, FFBEF04, FFTOT05)

2.6.4.1 Definition of Flat Fee Payments

A flat fee is the fixed dollar amount a person is charged for a package of health care services provided during a defined period of time. Examples would be: obstetrician's fee covering a normal delivery, as well as pre- and post-natal care; or a surgeon's fee covering surgical procedure and post-surgical care. A flat fee group is the set of medical services (i.e., events) that are covered under the same flat fee payment. The flat fee groups represented on the STAZ file include flat fee groups where at least one of the health care events, as reported by the HC respondent, occurred during 2004. By definition, a flat fee group can span multiple years. Furthermore, a single person can have multiple flat fee groups.

2.6.4.2 Flat Fee Variable Descriptions

2.6.4.2.1 Flat Fee ID (FFEEIDX)

As noted earlier in Section 2.6.1.2 "Record Identifiers," the variable FFEEIDX uniquely identifies all events that are part of the same flat fee group for a person. On any 2004 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.

2.6.4.2.2 Flat Fee Type (FFIPTYPE)

FFIPTYPE indicates whether the 2004 hospital stay is the "stem" or "leaf" of a flat fee group. A stem (records with FFIPTYPE = 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 FFIPTYPE = 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 hospital inpatient stays that are not part of a flat fee payment, the FFIPTYPE is set to –1, "INAPPLICABLE."

2.6.4.2.3 Counts of Flat Fee Events that Cross Years (FFBEF04, FFTOT05)

As explained in Section 2.6.4.1, a flat fee payment covers multiple events and the multiple events could span multiple years. For situations where the hospital inpatient stay/event occurred in 2004 as a part of a group of events, and some event occurred before or after 2004, counts of the known events are provided on the STAZ record. Variables that indicate events occurred before or after 2004 are as follows:

> FFBEF04 - total number of pre-2004 events in the same flat fee group as the 2004 hospital inpatient stay(s). This count would not include 2004 hospital

inpatient stay(s). Because there were no 2003 events for any flat fee group, this variable was omitted from the 2004 IP file.

FFTOT05 - the number of 2005 hospital inpatient stays expected to be in the same flat fee group as the hospital inpatient stay that occurred in 2004. Because there were no 2005 events expected for any flat fee group, this variable was omitted from the 2004 IP file.

If there are no 2003 events on the file, FFBEF04 will be omitted. Likewise, if there are no 2005 events on the file, FFTOT05 will be omitted. If there are no flat fee data related to the records in this file, FFEEIDX and FFIPTYPE will be omitted as well. Please note that the crosswalk in this document lists all possible flat fee variables.

2.6.4.3 Caveats of Flat Fee Groups

There is 1 hospital inpatient stays/events that are identified as being part of a flat fee payment group. 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 2004, but the remaining visits that were part of this flat fee group occurred in 2005. In this case, the 2004 flat fee group would consist of one event, the stem. The 2005 events that are part of this flat fee group are not represented on the file. Similarly, the household respondent may have reported a flat fee group where the initial visit began in 2003 but subsequent visits occurred during 2004. In this case, the initial visit would not be represented on the file. This 2004 flat fee group would then only consist of one or more leaf records and no stem.

2.6.5 Expenditure Data

2.6.5.1 Definition of Expenditures

Expenditure variables 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 for each hospital stay, 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 1990s due to the increasingly common practice of discounting. Although measuring expenditures as the sum of payments incorporates discounts in the MEPS expenditure estimates, these estimates do not incorporate any payment not directly tied to specific medical care 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 expenditures defined in the 1987 NMES. For details on expenditure definitions, please reference the following, "Informing American Health Care Policy" (Monheit, et al., 1999). AHRQ has developed factors to apply to the 1987 NMES expenditure data to facilitate longitudinal analysis. These factors can be assessed via the CFACT data center. For more information, see the Data Center section of the MEPS web site http://www.meps.ahrq.gov.

Expenditure data related to hospital inpatient events are broken out by facility and separately billing doctor expenditures. This file contains six categories of expenditure variables per stay: basic hospital facility expenses; expenses for doctors who billed separately from the hospital for any inpatient services provided during hospital stay; total expenses, which is the sum of the facility and physician expenses; facility charge; physician charge; and total charges, which is the sum of the facility and physician charges. If examining trends in MEPS expenditures or performing longitudinal analysis on MEPS expenditures, please refer to section C, sub-section 2.1 for more information.

2.6.5.2 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 Components (MPC). The MPC contacted medical providers identified by household respondents. The charge and payment data from medical providers were used in the expenditure imputation process to supplement missing household data. For all hospital inpatient stays, MPC data were used if available; otherwise, HC data were used. Missing data for hospital inpatient stays where HC data were not complete and MPC data were not collected, or MPC data were not complete, were imputed during the imputation process.

2.6.5.2.1 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, copayments or charges reported as total payments, and reimbursed amounts that were reported as out-of-pocket payments. In addition, edits were implemented to correct for misclassifications 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.

2.6.5.2.2 General Hot-Deck Imputation

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

2.6.5.2.3 Hospital Inpatient Stay Data Editing and Imputation

Facility expenditures for hospital inpatient stays 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 a 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. Most hospital inpatient stays were imputed as simple events because facility charges for an inpatient hospital stay are rarely grouped with other events. (See Section 2.6.4 for more details on flat fee groups.)

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 with missing expenditure data was assigned to a recipient category based on the extent of its missing charge and expenditure data. For example, an event with a known total charge but no expenditure information was assigned to one category, while an event with a known total charge and partial expenditure information was assigned to a different category. Similarly, events without a

known total charge and no or partial expenditure information were assigned to various recipient categories.

The logical edits produced eight recipient categories in which all events had a common extent of missing data. Separate hot-deck imputations were performed on events in each recipient category. For hospital inpatient and emergency room events, the donor pool was restricted to events with complete expenditures from the MPC. Due to the low ratio of donors to recipients for hospital outpatient and office based events, there were no donor pool restrictions.

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 distribution of free events among complete events (donors) would not be represented among incomplete events (recipients).

Expenditures for services provided by separately billing doctors in hospital settings were also edited and imputed. These expenditures are shown separately from hospital facility charges for hospital inpatient, outpatient, and emergency room care.

2.6.5.3 Imputation Flag (IMPFLAG)

IMPFLAG is a six-category variable that indicates if the event contains complete Household Component (HC) or Medical Provider Component (MPC) data, was fully or partially imputed, or was imputed in the capitated imputation process (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 flat fee leaf events)

IMPFLAG=1 complete HC data

IMPFLAG=2 complete MPC data

IMPFLAG=3 fully imputed

IMPFLAG=4 partially imputed

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

2.6.5.4 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 facility payments, while physician's expenditures may be still present. Thus, if the first visit in the flat fee group occurred prior to 2004, all of the events that occurred in 2004 will have zero payments. Conversely, if the first event in the flat fee group occurred at the end of 2004, the total expenditure for the entire flat fee group will be on that event, regardless of the number of events it covered after 2004. See Section 2.6.4 for details on the flat fee variables.

2.6.5.5 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, or (3) care for these events was covered under a flat fee arrangement with all expenditures appearing on the stem event. 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.

2.6.5.6 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.

2.6.5.7 Mother/Newborn Expenditures

Expenditure data for newborns were edited to exclude discharges after birth when the newborn left the hospital before or on the same day as the mother. As a result, inpatient expenditures reported for 2004 births were usually applied to the mother and not treated as separate expenditures for the infant. However, if a newborn was discharged at a later date than the mother's discharge date, then the hospitalization was treated as a separate hospital stay for the newborn.

This means that, in most cases, expenditure data for the newborn is included on the mother's record. A separate record for the newborn only exists if the newborn was discharged after the mother. In this case, the expenditure for the newborn is on the newborn's record.

2.6.5.8 Hospital Inpatient Stay/Emergency Room Expenditures

Although a person may have indicated that there was an emergency room visit that preceded this hospital stay (EMERROOM), there was no verification that, in fact, the emergency room visit was actually recorded within the Emergency Room Section of the questionnaire.

While it is true that all of the event files can be linked by DUPERSID, there is no unique record link between hospital inpatient stays and emergency room visits. That is, a person could have one hospital inpatient stay and three emergency room visits during the calendar year. While the hospital inpatient stay record may indicate that it was preceded by an emergency room visit, there is no unique record link to the appropriate (of the three) emergency room visit.

However, wherever relationships could be identified (using the MPC start and end date of the events as well as other information from the provider), the facility expenditure associated with the emergency room visit was moved to the hospital facility expenditure. Hence, for some hospital stays, facility expenditures for a preceding emergency room visit are included. In these situations, the corresponding emergency room record on the MEPS 2004 Emergency Room Visit File will have its facility expenditure information zeroed out to avoid double-counting. The variable ERHEVIDX identifies these hospital stays whose expenditures include the facility expenditures for the preceding emergency room visit (see ERHEVIDX in Section 2.6.1.2). It should also be noted that for these cases, there is only one hospital stay associated with the emergency room stay.

2.6.5.9 Sources of Payment

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

1. Out-of-pocket by user or family,
2. Medicare,
3. Medicaid,
4. Private Insurance,
5. Veterans Administration, excluding TRICARE,
6. TRICARE,
7. Other Federal sources - includes Indian Health Service, Military Treatment Facilities, and other care by the Federal government,
8. Other State and Local Source - includes community and neighborhood clinics, State and local health departments, and State programs other than Medicaid,
9. Workers' Compensation, and

10. Other Unclassified Sources - includes sources such as automobile, homeowner's, and liability insurance, and other miscellaneous or unknown sources.

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

2.6.5.10 Imputed Hospital Inpatient Stay Expenditure Variables

This file contains two sets of imputed expenditure variables: facility expenditures and physician expenditures.

2.6.5.10.1 Hospital Inpatient Facility Expenditures (IPFSF04X-IPFOT04X, IPFXP04X, IPFTC04X)

Hospital facility expenses include all expenses for direct hospital care, including room and board, diagnostic and laboratory work, x-rays, and similar charges, as well as any physician services included in the hospital charge.

IPFSF04X - IPFOT04X are the 12 sources of payment. The 12 sources of payment are: self/family (IPFSF04X), Medicare (IPFMR04X), Medicaid (IPFMD04X), private insurance (IPFPV04X), Veterans Administration (IPFVA04X), TRICARE (IPFTR04X), other Federal sources (IPFOF04X), State and Local (non-federal) government sources

(IPFSL04X), Worker's Compensation (IPFWC04X), other private insurance (IPFOR04X), other public insurance (IPFOU04X), and other insurance (IPFOT04X). IPFXP04X is the sum of the 12 sources of payment for the Hospital Facility expenditures, and IPFTC04X is the total charge.

Wherever an emergency room visit record is linked to a hospital inpatient stays record (identified by the variable ERHEVIDX, see Section 2.6.1.2), the facility source of payment variables on the emergency room visit record were zeroed out because the emergency room expenditures were already included in the hospital facility source of payment variables.

2.6.5.10.2 Hospital Inpatient Physician Expenditures (IPDSF04X - IPDOT04X, IPDTC04X, IPDXP04X)

Separately billing doctor (SBD) expenses typically cover services provided to patients in hospital settings by providers like anesthesiologists, radiologists, and pathologists, whose charges are often not included in hospital bills.

For medical doctors who bill separately (i.e., outside the hospital bill), a separate data collection effort within the Medical Provider Component was performed to obtain this same set of expenditure information from each separately billing doctor. It should be noted that there could be several separately billing doctors associated with a medical event. For example, a hospital inpatient stay could have a radiologist, anesthesiologist, pathologist and a surgeon associated with it. If their services are not included in the hospital bill then this is one medical event with four separately billing doctors. The imputed expenditure information associated with the separately billing doctors for a hospital inpatient stay is combined (i.e., the expenditures incurred by the radiologist + anesthesiologist + pathologist + surgeon) and is provided on the file. IPDSF04X - IPDOT04X are the 12 sources of payment; IPDXP04X is the sum of the 12 sources of payments; and IPDTC04X is the physician's total charge.

Data users/analysts need to take into consideration whether to analyze facility and SBD expenditures separately, combine them within service categories, or collapse them across service categories (e.g., combine SBD expenditures with expenditures for physician visits to offices and/or outpatient departments).

2.6.5.10.3 Total Expenditures and Charges for Hospital Inpatient Stays (IPXP04X and IPTC04X)

Data users/analysts interested in total expenditures should use the variable IPXP04X, which includes both facility and physician amounts. Those interested in total charges should use the variable IPTC04X, which includes both facility and physician charges (see Section 2.6.5.1 for an explanation of the "charge" concept).

2.6.5.11 Rounding

Expenditure variables have been rounded to the nearest penny. Person-level expenditure information released on the MEPS 2004 Person-Level Use and Expenditure File were rounded to the nearest dollar. It should be noted that using the MEPS 2004 event files to create person-level totals will yield slightly different totals than 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 MEPS 2004 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 2004 Appendix File, HC-085I, for details on such rounding differences.

3.0 Sample Weight (PERWT04F)

3.1 Overview

There is a single full year person-level weight (PERWT04F) 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 2004. A key person either was a member of an NHIS household at the time of the NHIS interview, or became a member of a family associated with such a household after being out-of-scope at the time of the NHIS (the latter circumstance includes newborns as well as 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 PERWT04F was developed in several stages. Person-level weights for Panels 8 and 9 were created separately. The weighting process for each panel included an adjustment for nonresponse over time and calibration to independent population figures. The calibration was initially accomplished separately for each panel by raking the corresponding sample weights to Current Population Survey (CPS) population estimates based on five variables. The five variables used in the establishment of the initial person-level control figures were: census region (Northeast, Midwest, South, West); MSA status (MSA, non-MSA); race/ethnicity (Hispanic, non-Hispanic with black as sole reported race, non-Hispanic with Asian as sole reported race, and other); sex; and age. A 2004 composite weight was then formed by multiplying each weight from Panel 8 by the factor .49 and each weight from Panel 9 by the factor .51. The choice of factors reflected the relative sample sizes of the two panels, helping to limit the variance of estimates obtained from pooling the two samples. The composite weight was again raked to the same set of CPS-based control totals. When poverty status information derived from income variables became available, a final raking was undertaken on the previously established weight variable. Control totals were established using poverty status (five

categories: below poverty, from 100 to 125 percent of poverty, from 125 to 200 percent of poverty, from 200 to 400 percent of poverty, at least 400 percent of poverty) as well as the original five variables used in the previous calibrations.

3.2.1 MEPS Panel 8 Weight

The person-level weight for MEPS Panel 8 was developed using the 2003 full year weight for an individual as a "base" weight for survey participants present in 2003. For key, in-scope respondents who joined an RU some time in 2004 after being out-of-scope in 2003, the 2003 family weight associated with the family the person joined served as a "base" weight. The weighting process included an adjustment for nonresponse over Rounds 4 and 5 as well as raking to population control figures for December 2004. These control figures were derived by scaling back the population totals obtained from the March 2004 CPS to correspond to a national estimate for the civilian noninstitutionalized population provided by the Census Bureau for December 2004. Variables used in the establishment of person-level control figures included: census region (Northeast, Midwest, South, West); MSA status (MSA, non-MSA); race/ethnicity (Hispanic, black but non-Hispanic, Asian but non-Hispanic, and other); sex; and age. Overall, the weighted population estimate for the civilian noninstitutionalized population on December 31, 2004 is 289,659,890. Key, responding persons not in-scope on December 31, 2004 but in-scope earlier in the year retained, as their final Panel 8 weight, the weight after the nonresponse adjustment.

3.2.2 MEPS Panel 9 Weight

The person-level weight for MEPS Panel 9 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 2004 portion of Round 3 as well as raking to the same population control figures for December 2004 used for the MEPS Panel 8 weights. The same five variables employed for Panel 8 raking (census region, MSA status, race/ethnicity, sex, and age) were used for Panel 9 raking. Similarly, for Panel 9, key, responding persons not in-scope on December 31, 2004 but in-scope earlier in the year retained, as their final Panel 9 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 2004 CPS data base.

3.2.3 The Final Weight for 2004

Variables used in the establishment of person-level control figures included: poverty status (below poverty, from 100 to 125 percent of poverty, from 125 to 200 percent of poverty, from 200 to 400 percent of poverty, at least 400 percent of poverty); census region (Northeast, Midwest, South, West); MSA status (MSA, non-MSA); race/ethnicity (Hispanic, non-Hispanic with black as sole reported race, non-Hispanic with Asian as sole reported race, and other); sex; and age. Overall, the weighted population estimate for the civilian noninstitutionalized population for December 31, 2004 is 289,659,890 (PERWT04F>0 and INSC1231=1). The weights of some persons out-of-scope on December 31, 2004 were also calibrated, this time using poststratification. Specifically, the weights of persons out-of-scope on December 31, 2004 who were in-scope some time during the year and also entered a nursing home during the year were poststratified to a corresponding control total obtained from the 1996 MEPS Nursing Home Component. The weights of persons who died while in-scope during 2004 were poststratified to corresponding estimates derived using data obtained from the Medicare Current Beneficiary Survey (MCBS) and Vital Statistics information provided by the National Center for Health Statistics (NCHS). Separate control totals were developed for the "65 and older" and "under 65" civilian noninstitutionalized populations. The sum of the person-level weights across all persons assigned a positive person level weight is 293,527,003.

3.2.4 Coverage

The target population for MEPS in this file is the 2004 U.S. civilian noninstitutionalized population. However, the MEPS sampled households are a subsample of the NHIS households interviewed in 2002 (Panel 8) and 2003 (Panel 9). New households created after the NHIS interviews for the respective Panels and consisting exclusively of persons who entered the target population after 2002 (Panel 8) or after 2003 (Panel 9) are not covered by MEPS. Neither are previously out-of-scope persons who join an existing household but are unrelated to the current household residents. Persons not covered by a given MEPS panel thus include some members of the following groups: immigrants; persons leaving the military; U.S. citizens returning from residence in another country; and persons leaving institutions. The set of uncovered persons constitutes only a small segment of the MEPS target population.

4.0 Strategies for Estimation

This file is constructed for efficient estimation of utilization, expenditures, and sources of payment for hospital inpatient care and to allow for estimates of number of persons with hospital inpatient utilization for 2004.

4.1 Variables with Missing Values

It is essential that the data user/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 data user/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 data user/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 expenditure) are described in Section 2.6.5.

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 hospital inpatient stays, expenditures and sources of payment, the value in each record contributing to the estimates must be multiplied by the weight (PERWT04F) contained on that record.

Example 1

For example, the total number of hospital inpatient stays, regardless of the length of the hospital stay, for the civilian noninstitutionalized population of the U.S. in 2004, is estimated as the sum of the weight (PERWT04F) across all records. That is,

$$\Sigma W_j = 30{,}194{,}246 \qquad (1)$$

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 at the hospital inpatient stay level (where the visit has a total expense greater than 0), should be calculated as the weighted mean of the facility bill and doctor's bill paid by self/family. That is,

$$(\sum W_j X_j)/(\sum W_j) = \$225.91 \qquad (2)$$

where $X_j = IPFSF04X_j + IPDSF04X_j$ and $\sum W_j = 29{,}800{,}382$

for all records with $IPXP04X_j > 0$.

This gives $225.91 as the estimated mean amount of out-of-pocket payment of expenditures associated with hospital inpatient stays (discharges) and 29,800,382 as an estimate of the total number of hospital inpatient stays with expenditures. Both of these estimates are for the civilian noninstitutionalized population of the U.S. in 2004.

Example 3

Another example would be to estimate the mean proportion of total expenditures paid by private insurance for hospital inpatient stays with expenditures. This should be calculated as the weighted mean of the proportion of total expenditures paid by private insurance at the event level. That is,

$$(\sum W_j Y_j)/(\sum W_j) = 0.3879 \qquad (3)$$

where $Y_j = (IPFPV04X_j + IPDPV04X_j)/IPXP04X_j$ and $\sum W_j = 29{,}800{,}382$

for all records with $IPXP04X_j > 0$.

This gives 0.3879 as the estimated mean proportion of total expenditures paid by private insurance for hospital inpatient stays with expenditures for the civilian noninstitutionalized population of the U.S. in 2004.

4.3 Estimates of the Number of Persons with Hospital Inpatient Stays

When calculating an estimate of the total number of persons with hospital inpatient stays, users can use a person-level file or this event file. However, this event file must be used when the measure of interest is defined at the event level. For example, to estimate the number of persons in the civilian noninstitutionalized population of the U.S. discharged from a hospital in 2004 with at least one hospital stay of 10 or more nights, this event file must be used. This would be estimated as,

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

where

W_i is the sampling weight (PERWT04F) for person i

and

X_i = 1 if NUMNIGHX$_j$ GE 10 for any stay of person i
= 0 otherwise.

4.4 Person-Based Ratio Estimates

4.4.1 Person-Based Ratio Estimates Relative to Persons with Hospital Inpatient Use

This file may be used to derive person-based ratio estimates. However, when calculating ratio estimates where the denominator is at the person level, care should be taken to properly define and estimate the unit of analysis as person-level. For example, the mean expense for persons with hospital inpatient stays is estimated as:

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

where

W_i is the sampling weight (PERWT04F) for person i

and

Z_i = $\sum$ IPXP04X$_j$ across all stays 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 hospital inpatient stay 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 use and those without use). For example, the proportion of the civilian noninstitutionalized population of the U.S. with at least one hospital inpatient stay of four or more days would be estimated as:

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

where

W_i is the sampling weight (PERWT04F) for person i

and

Z_i = 1 if NUMNIGHX$_j$ GE 4 for any stay 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 with another, the major analytical variable (i.e., the dependent variable) determines the correct sampling weight to use.

4.6 Variance Estimation (VARSTR, VARPSU)

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 2004 data. Variables needed to implement a Taylor Series estimation approach are provided in the file and are described in the paragraph below.

Using a Taylor Series approach, variance estimation strata and the variance estimation PSUs within these strata must be specified. The corresponding variables on the MEPS full year utilization database are VARSTR and VARPSU, respectively. Prior to 2004, MEPS variance strata and PSUs were developed independently from year to year, and the last two characters of the strata and PSU variable names denoted the year. However, beginning with the 2004 Point-in-Time PUF, the variance strata and PSUs have been developed to be compatible with all future PUFs. Thus, data from future years can be pooled and the variance strata and PSU variables provided can be used without modification for variance estimation purposes for estimates covering multiple years of data. There are 203 variance estimation strata, each stratum with either two or three variance estimation PSUs. Specifying a "with replacement" design in a computer software package such as SUDAAN (Shah, et al, 1996) should provide standard errors appropriate for assessing the variability of MEPS survey estimates. It should be noted that the number of degrees of freedom associated with estimates of variability indicated by such a package may not appropriately reflect the actual number available. For MEPS sample estimates for characteristics generally distributed throughout the country (and thus the sample PSUs), there are over 100 degrees of freedom associated with the corresponding estimates of variance. The following illustrates these concepts using two examples from section 4.2.

Examples 2 and 3 from Section 4.2

Using a Taylor Series approach, specifying VARSTR and VARPSU as the variance estimation strata and PSUs (within these strata), respectively, and specifying a "with replacement" design in a computer software package (i.e., SUDAAN) will yield standard error estimates of $18.92 and 0.0131 for the estimated mean 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 the MEPS 2004 Hospital Inpatient Stays File can be used alone or in conjunction with other files. This section provides instructions for linking the hospital stays file with other MEPS public use files, namely, the person-level file, the prescribed medicines file, and the medical conditions file.

5.1 Merging a 2004 Person-Level File to the 2004 Hospital Inpatient Stays File

Merging characteristics of interest from other MEPS files (e.g., MEPS 2004 Full-Year Population Characteristics File) expands the scope of potential estimates. For example, to estimate the total number of hospital inpatient stays for persons with specific demographic characteristics (such as, age, race, sex, and education), population characteristics from a person-level file need to be merged onto the hospital inpatient stays file. This procedure is illustrated below. The MEPS 2004 Appendix File, HC-085I, provides additional detail on how to merge MEPS data files.

1. Create data set PERSX by sorting the MEPS 2004 Full Year Population Characteristics File by the person identifier, DUPERSID. Keep only variables to be merged onto the hospital inpatient stays file, and DUPERSID.

2. Create data set STAZ by sorting the hospital inpatient stays file by person identifier, DUPERSID.

3. Create final data set NEWSTAZ by merging these two files by DUPERSID, keeping only records on the hospital inpatient stays file.

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

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

```
RUN;

PROC SORT DATA=STAZ;
   BY DUPERSID;
RUN;

DATA NEWSTAZ;
   MERGE STAZ (IN=A) PERSX(IN=B);
   BY DUPERSID;
   IF A;
RUN;
```

5.2 Linking the 2004 Hospital Inpatient Stays File to the 2004 Medical Conditions File and/or the 2004 Prescribed Medicines File

Due to survey design issues, data users/analysts must keep limitations and caveats in mind when linking the different files. Those limitations/caveats are listed below. For detailed linking examples, including SAS code, data users/analysts should refer to the MEPS 2004 Appendix File, HC-085I.

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

The RXLK file provides a link from the MEPS event files to records on the 2004 Prescribed Medicine File. When using RXLK, data users/analysts should keep in mind that one hospital inpatient stay could link to more than one prescribed medicine record. Conversely, a prescribed medicine event may link to more than one hospital inpatient stay or different types of events. When this occurs, it is up to the data user/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 2004 Medical Conditions File. When using the CLNK, data users/analysts should keep in mind that (1) conditions are self-reported, (2) there may be multiple conditions associated with a hospital inpatient stay, and (3) a condition may link to more than one hospital inpatient stay or any other type of visit. Data users/analysts should also note that not all hospital inpatient stays link to the medical conditions file.

References

Cohen, S.B. (1999). Sample Design of the 1996 Medical Expenditure Panel Survey Medical Provider Component. Journal of Economic and Social Measurement. Vol. 24, 25-53.

Cohen, S.B. (1997). Sample Design of the 1996 Medical Expenditure Panel Survey Household Component. Rockville (MD): Agency for Health Care Policy and Research; 1997. *MEPS Methodology Report, No. 2.* AHCPR Pub. No. 97-0027.

Cohen, J.W. (1997). Design and Methods of the Medical Expenditure Panel Survey Household Component. Rockville (MD): Agency for Health Care Policy and Research; 1997. *MEPS Methodology Report, No. 1.* AHCPR Pub. No. 97-0026.

Cohen, S.B. (1996). The Redesign of the Medical Expenditure Panel Survey: A Component of the DHHS Survey Integration Plan. *Proceedings of the COPAFS Seminar on Statistical Methodology in the Public Service.*

Cox, B.G. and Cohen, S.B. (1985). Chapter 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: Health 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-085D: 2004 HOSPITAL INPATIENT STAYS

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
ERHEVIDX	Event ID for corresponding emergency room visit	Constructed
FFEEIDX	Flat fee ID	CAPI derived
MPCDATA	MPC Data Flag	Constructed

Characteristics of Hospital Inpatient Stays

Variable	Description	Source
IPBEGYR	Event start date – year	CAPI derived
IPBEGMM	Event start date – month	CAPI derived
IPBEGDD	Event start date – day	CAPI derived
IPENDYR	Event end date – year	CAPI derived
IPENDMM	Event end date – month	CAPI derived
IPENDDD	Event end date – day	CAPI derived
NUMNIGHX	# of nights in hospital - Edited/Imputed	(Edited/Imputed)
NUMNIGHT	Number of nights stayed at provider	HS01
EMERROOM	Did stay begin with emergency room visit	HS02
SPECCOND	Hospital stay related to condition	HS03
RSNINHOS	Reason entered hospital	HS05
ANYOPER	Any operations or surgeries performed	HS06

Variable	Description	Source
VAPLACE	VA facility flag	Constructed
IPICD1X	3 digit ICD-9-CM condition code	Edited
IPICD2X	3 digit ICD-9-CM condition code	Edited
IPICD3X	3 digit ICD-9-CM condition code	Edited
IPICD4X	3 digit ICD-9-CM condition code	Edited
IPPRO1X	2 digit ICD-9-CM procedure code	Edited
IPPRO2X	2 digit ICD-9-CM procedure code	Edited
IPCCC1X	Modified Clinical Classification Code	Constructed/Edited
IPCCC2X	Modified Clinical Classification Code	Constructed/Edited
IPCCC3X	Modified Clinical Classification Code	Constructed/Edited
IPCCC4X	Modified Clinical Classification Code	Constructed/Edited
DSCHPMED	Medicines prescribed at discharge	HS08

Flat Fee Variables

Variable	Description	Source
FFIPTYPE	Flat Fee Bundle	Constructed
FFBEF04	Total # of visits in FF before 2004	FF05
FFTOT05	Total # of visits in FF after 2004	FF10

Imputed Total Expenditure Variables

Variable	Description	Source
IPXP04X	Total expenditure for event (IPFXP04X+IPDXP04X)	Constructed
IPTC04X	Total charge for event (IPFTC04X+IPDTC04X)	Constructed

Imputed Facility Expenditure Variables

Variable	Description	Source
IPFSF04X	Facility amount paid, self/family (Imputed)	CP Section (Edited)
IPFMR04X	Facility amount paid, Medicare (Imputed)	CP Section (Edited)
IPFMD04X	Facility amount paid, Medicaid (Imputed)	CP Section (Edited)
IPFPV04X	Facility amount paid, private insurance (Imputed)	CP Section (Edited)
IPFVA04X	Facility amount paid, Veterans Administration (Imputed)	CP Section (Edited)
IPFTR04X	Facility amount paid, TRICARE (Imputed)	CP Section (Edited)
IPFOF04X	Facility amount paid, other federal (Imputed)	CP Section (Edited)
IPFSL04X	Facility amount paid state & local government (Imputed)	CP Section (Edited)
IPFWC04X	Facility amount paid, workers' compensation (Imputed)	CP Section (Edited)
IPFOR04X	Facility amount paid, other private (Imputed)	Constructed
IPFOU04X	Facility amount paid, other pub (Imputed)	Constructed
IPFOT04X	Facility amount paid, other insurance (Imputed)	CP Section (Edited)
IPFXP04X	Facility sum payments IPFSF04X – IPFOT04X	Constructed
IPFTC04X	Total facility charge (Imputed)	CP Section (Edited)

Imputed Separately Billing Physician Expenditure Variables

Variable	Description	Source
IPDSF04X	Doctor amount paid, family (Imputed)	Constructed
IPDMR04X	Doctor amount paid, Medicare (Imputed)	Constructed
IPDMD04X	Doctor amount paid, Medicaid (Imputed)	Constructed
IPDPV04X	Doctor amount paid, private insurance (Imputed)	Constructed
IPDVA04X	Doctor amount paid, Veterans Administration (Imputed)	Constructed
IPDTR04X	Doctor amount paid, TRICARE (Imputed)	Constructed
IPDOF04X	Doctor amount paid, other federal (Imputed)	Constructed

Variable	Description	Source
IPDSL04X	Doctor amount paid, state & local government (Imputed)	Constructed
IPDWC04X	Doctor amount paid, workers' compensation (Imputed)	Constructed
IPDOR04X	Doctor amount paid, other private insurance (Imputed)	Constructed
IPDOU04X	Doctor amount paid, other public insurance (Imputed)	Constructed
IPDOT04X	Doctor amount paid, other insurance (Imputed)	Constructed
IPDXP04X	Doctor sum payments IPDSF04X–IPDOT04X	Constructed
IPDTC04X	Total doctor charge (Imputed)	Constructed
IMPFLAG	Imputation status	Constructed

Weights

Variable	Description	Source
PERWT04F	Expenditure file person weight, 2004	Constructed
VARSTR	Variance estimation stratum, 2004	Constructed
VARPSU	Variance estimation PSU, 2004	Constructed