MEPS HC 239H:
2022 Home Health Visits

May 2024

Agency for Healthcare Research and Quality
Center for Financing, Access, and Cost Trends
5600 Fishers Lane
Rockville, MD 20857
(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 dataset 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 dataset with individually identifiable records from any datasets other than the Medical Expenditure Panel Survey or the National Health Interview Survey. Furthermore, linkage of the Medical Expenditure Panel Survey and the National Health Interview Survey may not occur outside the AHRQ Data Center, NCHS Research Data Center (RDC) or the U.S. Census RDC network.

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

1.0 Household Component

The Medical Expenditure Panel Survey (MEPS) provides nationally representative estimates of health care use, expenditures, sources of payment, and health insurance coverage for the U.S. civilian noninstitutionalized population. The MEPS Household Component (HC) also provides estimates of respondents' health status, demographic and socio-economic characteristics, employment, access to care, and satisfaction with care. Estimates can be produced for individuals, families, and selected population subgroups. The panel design of the survey includes five rounds of interviews covering 2 full calendar years. Additional rounds were added to Panel 24 in 2021 and 2022, covering the third and fourth years, respectively, to compensate for the smaller number of completed interviews in later panels. These extra rounds provide data for examining person-level changes in selected variables such as expenditures, health insurance coverage, and health status. Information about each household member is collected through computer assisted personal interviewing (CAPI) technology and the survey builds on this information from interview to interview. All data for a sampled household are reported by a single household respondent.

The MEPS HC was initiated in 1996. Each year, a new panel of sample households is selected. Because the data collected are comparable to those from earlier medical expenditure surveys conducted in 1977 and 1987, it is possible to analyze long-term trends. Historically, each annual MEPS HC sample consists of approximately up to 15,000 households. Data can be analyzed at the person, the family, or the event level. Data must be weighted to produce national estimates.

The set of households selected for each panel of the MEPS HC is a subsample of households participating in the previous year's National Health Interview Survey (NHIS) conducted by the National Center for Health Statistics (NCHS). The NHIS sampling frame provides a nationally representative sample of the U.S. civilian noninstitutionalized population. In 2006, the NCHS implemented a new sample design for the NHIS to include households with Asian persons in addition to households with Black and Hispanic persons in the oversampling of minority populations. In 2016, NCHS introduced another sample design that discontinued the oversampling of these minority groups.

2.0 Medical Provider Component

When the household CAPI interview is completed, and permission is obtained from the household survey respondents to contact their medical provider(s), a sample of these providers is contacted by telephone to obtain information that household respondents cannot accurately provide. This part of the MEPS is called the Medical Provider Component (MPC), and it collects information on dates of visits, diagnosis and procedure codes, and charges and payments. The Pharmacy Component (PC), a subcomponent of the MPC, does not collect data on charges or on diagnosis and procedure codes, but it does collect detailed information on drugs, including the National Drug Code (NDC) and medicine name, as well as amounts of payment. The MPC is not

designed to yield national estimates. It is primarily used as an imputation source to supplement/replace household reported expenditure information.

3.0 Survey Management and Data Collection

MEPS HC and MPC data are collected under the authority of the Public Health Service Act. The MEPS HC data are collected under contract with Westat, Inc. and the MEPS MPC data are collected under contract with Research Triangle Institute. Datasets and summary statistics are edited and published in accordance with the confidentiality provisions of the Public Health Service Act and the Privacy Act. The NCHS provides consultation and technical assistance.

As soon as the MEPS data are collected and edited, they are released to the public in stages of microdata files and tables via the MEPS website and datatools.ahrq.gov.

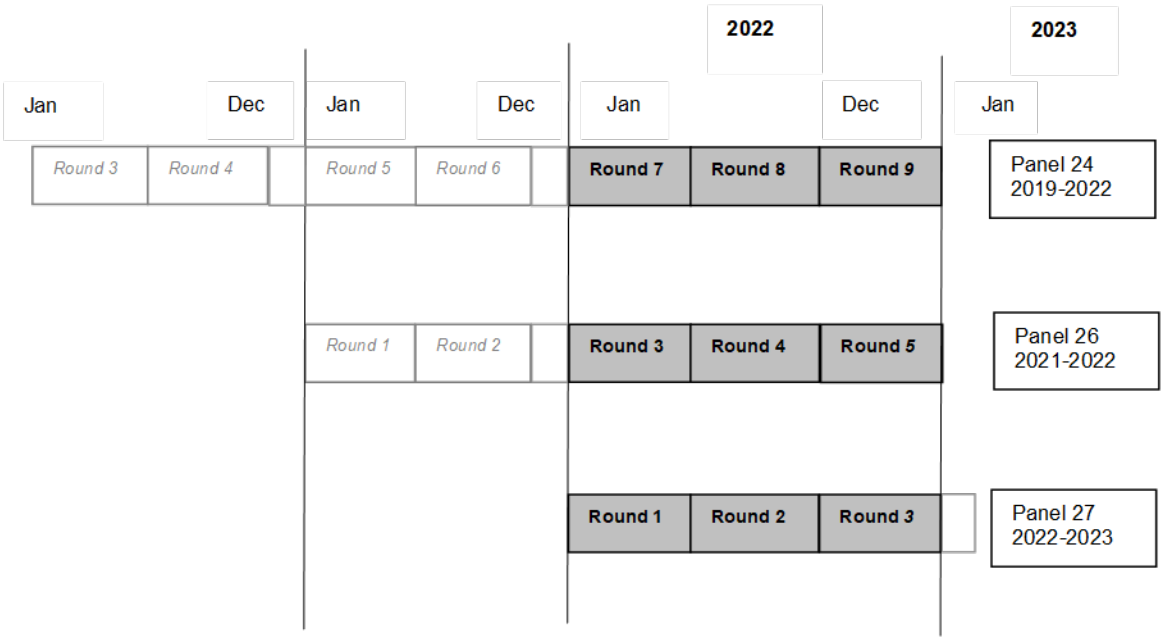
Additional information on MEPS is available from the MEPS project manager or the MEPS public use data manager at the Center for Financing, Access, and Cost Trends, Agency for Healthcare Research and Quality, 5600 Fishers Lane, Rockville, MD 20857 (301-427-1406).

C. Technical and Programming Information

1.0 General Information

This documentation describes one in a series of public use event files from the 2022 MEPS HC and MPC. It was released as an ASCII file (with related SAS, SPSS, R, and Stata programming statements and data user information) and as a SAS dataset, SAS transport file, Stata dataset, and Excel file. The 2022 Home Health Visits Public Use File (hereafter referred to as the HHV PUF) provides detailed information on home health events for a nationally representative sample of the U.S. civilian noninstitutionalized population. Data from the HHV PUF can be used to make estimates of home health (HH) event utilization and expenditures for the calendar year 2022. The PUF contains 51 variables and has a logical record length of 204 with an additional 2-byte carriage return/line feed at the end of each record. As illustrated below, this file consists of MEPS survey data obtained in the 2022 portion of Round 7 and all of Rounds 8 and 9 for Panel 24; the 2022 portion of Round 3 and all of Rounds 4 and 5 for Panel 26; and Rounds 1 and 2, and the 2022 portion of Round 3 for Panel 27 (i.e., the rounds for the MEPS panels covering the calendar year 2022).

Full year (FY) 2022 includes three panels of data; Panel 24 was extended to include Rounds 7, 8, and 9.



Counts of home health utilization are based entirely on household reports. Agency home health providers were sampled into the MEPS MPC (see Section B. 2.0). Only those providers for whom the respondent signed a permission form were included in the MPC. Information from the

MPC was used to supplement expenditure and payment data reported by the household, and does not affect use estimates.

Data from this event PUF can be merged with other 2022 MEPS HC PUFs to append person-level data such as demographic characteristics or health insurance coverage to each home health record.

This PUF can also be used to construct summary variables for expenditures, sources of payment, and related aspects of home health events for the calendar year 2022. Aggregate annual person-level information on the use of home health providers and other health services is provided on the Full Year 2022 Consolidated Public Use File (hereafter referred to as the Consolidated PUF), where each record represents a MEPS sampled person.

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

- Data File Information (Section 2.0)
- Survey Sample Information (Section 3.0)
- Strategies for Estimation (Section 4.0)
- Merging/Linking MEPS Data Files (Section 5.0)
- Variable-Source Crosswalk (Section D)

For more information on the MEPS HC sample design, see Chowdhury, et al. (2019). For information on the MEPS MPC design, see RTI (2019). A copy of the survey instrument used to collect the information in this PUF is available on the MEPS website.

2.0 Data File Information

The 2022 HHV PUF consists of one event-level file. The PUF contains characteristics associated with the home health event and imputed expenditure data.

The home health services represented in this PUF are provided by three kinds of home health providers: formal (paid) home health agency providers, paid independent providers (self-employed), and informal providers who do not reside in the same household as the MEPS sampled person (care from informal providers who live in the same household as the sampled person are not represented on this file).

Each record on this PUF represents a household-reported home health event. A home health event represents a MONTH of similar services provided to a sampled person by the same PROVIDER (i.e., an employer in the case of formal agency care and an individual in the case of paid independent and informal care providers). For example, if a person received from Provider Agency A four visits from a nurse, ten visits from a homemaker, and four visits from a physical therapist each month during the months of January, February, and March, and also received from

Provider B a physician visit in the months of January and February, there would be five event records on the file (NOT 56 records). There would be one event record representing all the visits from Provider A for the month of January, another record for Provider A February visits, a third Provider A record for the March visits, a fourth record representing Provider B physician visit in January, and a fifth representing the Provider B physician visit in February. Data were collected (and represented on this file) in this manner because agencies, hospitals, and nursing homes provide MEPS expenditure data in this manner. To be consistent with the definition of what is considered a home health event on this file, this same definition (i.e., a month of similar services) was applied to all types of home health providers.

The 2022 HHV PUF contains 7,540 home health records; of these records, 7,384 are associated with persons having a positive person-level weight (PERWT22F). It includes all records related to home health events for all household members who resided in eligible responding households and for whom at least one home health event was reported. Each record represents one household-reported home health event that occurred during the calendar year 2022. Some persons may have been reported to have multiple events and thus will be represented in multiple records on the PUF. Other persons may have no home health events reported and thus will have no records on this PUF. These data were collected during the 2022 portion of Round 7 and all of Rounds 8 and 9 for Panel 24; and the 2022 portion of Round 3, and all of Rounds 4 and 5 for Panel 26; and Rounds 1 and 2, the 2022 portion of Round 3 for Panel 27 of the MEPS HC. The persons represented in this PUF had to meet either (a) or (b) below:

a) Be classified as a Key in-scope person who responded for their entire period of 2022 eligibility (i.e., persons with a positive 2022 full-year person-level weight, PERWT22F > 0), or

b) Be an eligible member of a family whose Key in-scope members have a positive person-level weight (PERWT22F > 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 (FAMWT22F > 0). Note that FAMIDYR and FAMWT22F are variables on the 2022 Consolidated PUF.

Persons with no home health events for 2022 are not included on this event-level HHV PUF but are represented on the person-level 2022 Full-Year Population Characteristics PUF (hereafter referred to as the Population Characteristics PUF).

Home health providers include formal, i.e., paid, and informal, i.e., unpaid, providers. Formal or paid providers include home health agencies and other independent paid providers. Informal or unpaid providers include family and friends that reside outside of the sampled person's household.

For home health agencies it is important to distinguish between the provider and the home health worker. In these cases, the provider is the agency or the facility that employs the workers. The home health workers are the people who administer the care. Examples of home health care workers are the following: nurses, physical therapists, home health aides, homemakers, and hospice workers, among others. These examples are generally the types of workers associated with agencies. Paid independent providers generally include companions, nursing assistants,

physicians, etc. For each record on this file, one or more types of workers can be reported. The respondent is asked to mention all of the types of home health workers who provided home health care (since records represent a month of service, there can be more than one type of worker on a single record). For example, an agency that provides two types of aides that provide home health care to the same person during a specific month is represented as one event on the file even though two workers employed at the same agency provided care. When using this file, analysts must keep in mind that a record on the file corresponds to a provider entity, not an individual or particular worker.

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

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

To append person-level information such as demographic or health insurance coverage to each event record, data from this file can be merged with 2022 MEPS HC person-level data (e.g. Consolidated or Population Characteristics PUFs) using the person identifier, DUPERSID. Home health events can also be linked to the MEPS 2022 Medical Conditions file. Please see Section 5.0 or the Appendix to the MEPS 2022 Event Files (hereafter referred to as the Appendix PUF) for details on how to link MEPS data files.

2.1 Codebook Structure

For most variables on the HHV PUF, both weighted and unweighted frequencies are provided in the accompanying codebook. The exceptions to this are weight variables and variance estimation variables. Only unweighted frequencies of these variables are included in the accompanying codebook file. See the Weights Variables list in Section D, “Variable-Source Crosswalk”.

The codebook and data file list variables in the following order:

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

The person identifier corresponds to a unique person and the home health event identifier corresponds to a unique event.

2.2 Reserved Codes

This HHV PUF contains several reserved code values

Table 1

Reserved Code Values and Definitions

Value		Definition
-1	Inapplicable	Question was not asked due to skip pattern
-7	Refused	Question was asked and respondent refused to answer question
-8	Don't Know	Question was asked and respondent did not know answer or the information could not be ascertained
-15	Cannot Be Computed	Value cannot be derived from data

The value Cannot Be Computed (-15) is assigned to MEPS constructed variables in cases where there is not enough information from the MEPS instrument to calculate the constructed variables. "Not enough information" is often the result of skip patterns in the data or missing information stemming from the responses Refused (-7) or Don't Know (-8). Note that, in addition to Don't Know, reserved code -8 includes cases for which the information from the question was not ascertained.

Generally, values of -1, -7, -8, and -15 for non-expenditure variables have not been edited in this PUF. Analysts who would like to recode these values can find skip patterns in the HC survey questionnaire located on the MEPS website.

2.3 Codebook Format

The codebook describes an ASCII dataset (although the data are also being provided in a SAS dataset, a SAS transport file, a Stata dataset, and an Excel file) and provides programming identifiers for each variable.

Table 2

Programming Identifiers for Each Variable in the HHV PUF

Identifier	Description
Name	Variable name
Description	Variable descriptor

Identifier	Description
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, the variable names reflect the content of the variable. Generally, imputed/edited variables end with an "X".

As the collection, universe, or categories of variables were altered, some variable names have been appended with "_Myy", where "yy" indicates the collection year in which the alterations were made. Such alterations are described in detail throughout this document.

2.4.1 Variable-Source Crosswalk

Variables in this HHV PUF were derived either from the CAPI or the MPC data collection instrument. 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,"
2. Variables from one or more specific questions have those questionnaire sections and question numbers indicated in the "Source" column; questionnaire sections are identified as:
    - EV - Event Roster section
    - HH - Home Health Event section
    - CP - Charge Payment section
3. Variables constructed from multiple questions by 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 that they were edited/imputed. Please note that imputed means that a series of logical edits, as well as an imputation process to account for missing data, were performed on the variable.

The total sum of payments and the 10 source of payment variables are named in the following way:

The first two characters indicate the type of event:

IP - inpatient stay

ER - emergency room visit

HH - home health visit

OM - other medical equipment

OB - office-based visit

OP - outpatient visit

DV - dental visit

RX - prescribed medicine

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

SF - self or family

MR - Medicare

MD - Medicaid

PV - private insurance

VA - Veterans Administration/CHAMPVA

OF - other federal government

SL - state/local government

WC - Workers' Compensation

OT - other insurance

TR - TRICARE

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 (22). The seventh character, "X", indicates the variable is edited/imputed.

For example, HHSF22X is the edited/imputed amount paid by self or family for 2022 home health expenditures.

2.5 File Contents

2.5.1 Survey Administration Variables

Person Identifiers (DUID, PID, DUPERSID)

The definitions of Dwelling Units (DUs) in the MEPS Household Survey are generally consistent with the definitions employed for the NHIS. The dwelling unit ID (DUID) is a 7-digit number consisting of a 2-digit panel number followed by a 5-digit random number assigned after the case was sampled for MEPS. A 3-digit person number (PID) uniquely identifies each person within the DU. The variable DUPERSID is the combination of the variables DUID and PID. IDs begin with the 2-digit panel number.

For detailed information on dwelling units and families, please refer to the documentation for the 2022 Population Characteristics PUF.

Record Identifier (EVNTIDX)

EVNTIDX uniquely identifies each event (i.e., each record on this PUF) and is the variable required to link home health events to data files containing details on conditions (MEPS 2022 Medical Conditions PUF). EVNTIDX begins with the 2-digit panel number and ends with the 2-digit event type number. For details on linking see Section 5.0: Merging/Linking MEPS Data Files or the MEPS 2022 Appendix PUF, HC-239I.

Round Indicator (EVENTRN)

EVENTRN indicates the round in which the home health event was reported. Please note: Rounds 7 (partial), 8, and 9 are associated with MEPS data collected from Panel 24. Likewise, Rounds 3 (partial), 4, and 5 are associated with data collected from Panel 26; and Rounds 1, 2, and 3 (partial) are associated with data collected from Panel 27.

Panel Indicator (PANEL)

PANEL is a constructed variable used to specify the panel number for the person. PANEL will indicate either Panel 24, Panel 26, or Panel 27 for each person on the PUF. Panel 24 is the panel that started in 2019, Panel 26 is the panel that started in 2021, and Panel 27 is the panel that started in 2022.

2.5.2 Home Health Event Variables

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

Date of Event (HHDATEYR, HHDATEMM)

The date variables (HHDATEYR and HHDATEMM) indicate the year and month that the household respondent reported as the year and month of occurrence for this type of home health event. An artifact of the data collection for the variable HHDATEYR is that a person may have started receiving that type of home health care from that provider prior to 2022. These variables should not be interpreted as "true" start dates.

Characteristics of Event (MPCELIG-HCarWrkrNonProfNone_M18)

The HC questionnaire asked the respondent to indicate whether the home health provider event(s) for each month's services were provided through an agency or an independent paid provider (SELFAGEN). The response to the SELFAGEN question dictated the skip pattern CAPI followed regarding the questions in the home health section of the HC questionnaire. The questionnaire also asked respondents if the provider was paid or whether a friend, relative, or volunteer (HHTYPE) provided the home health services. The constructed variable MPCELIG indicates whether the home health provider event was eligible for MPC data collection and the type of imputation process the event went through. MPCELIG is a more accurate variable for determining whether the event was an agency, a paid independent, or an informal care event. However, SELFAGEN is a more accurate variable for determining the home health questions asked of the respondent. For all members receiving care from an agency, hospital, or nursing home, the respondent was asked to identify the type of skilled home health worker (CNA_M18-HCarWrkrProfNone_M18) and the type of non-skilled home health worker (COMPANN_M18-HCarWrkrNonProfNone_M18) they saw - for example, a certified nursing assistant as the skilled worker and a home health aide as the non-skilled worker.

Analysts should keep in mind that these identifications by household respondents are subjective in nature, are not mutually exclusive or collectively exhaustive, and should not be used to make certain estimates. For example, a person on one type of insurance may identify an individual providing home health care services to them as a personal care attendant while an individual having a different type of insurance coverage may identify that same worker as a home care aide. Making estimates of personal care attendants or home care aides based on their identification by household respondents and treating these types of workers as mutually exclusive groups will result in inaccurate estimates. Respondents may also have indicated that a person was seen by more than one home health care worker during a single event. For example, since an event is a month of services, a respondent may have reported that a person was seen by a nurse, a physical therapist, and/or a home health aide during a single event.

Frequency of Event and Visit Details (FREQCY-VSTRELCN)

Several variables identify the frequency and length of home health events (FREQCY-DAYSPMO) and whether or not the same services were received during each month (SAMESVCE). Frequency of event variables (FREQCY- DAYSPMO) were used as building blocks to construct HHDAYS. HHDAYS indicates the number of days the person received care during that event (i.e., month of care). Frequency variables can be combined to get a measure of

the intensity of care. Regardless of the type of provider, all respondents were asked if the home health services received were due to a medical condition (VSTRELCN).

2.5.3 Flat Fee Variables

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. Because MEPS does not collect flat fee information about home health events, no flat fee variables are included in this file.

2.5.4 Condition Codes

Information on household-reported medical conditions associated with each home health event is NOT provided on this PUF. To obtain complete condition information associated with an event, the analyst must link to the 2022 Medical Conditions PUF. Details on how to link to the 2022 Medical Conditions PUF are provided in the 2022 Appendix PUF, HC-239I.

2.5.5 Expenditure Data

Definition of Expenditures

Expenditures in this PUF refer to payments 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 from its predecessors, the 1987 NMES and 1977 NMCES surveys, where "charges" rather than the 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 events, such as bonuses or retrospective payment adjustments paid by third party payers. Another general change from the two prior surveys is that charges associated with uncollected liability, bad debt, and charitable care (unless provided by a public clinic or hospital) are not counted as expenditures because there are no associated payments. While charge data are provided in this PUF, analysts should use caution when working with these data because a charge does not typically represent actual dollars exchanged for services or the resource costs of those services, nor are they directly comparable to the expenditures defined in the 1987 NMES. For details on expenditure definitions, see Monheit, et al. (1999). AHRQ has developed factors to apply to the 1987 NMES expenditure data to facilitate longitudinal analysis. These factors are published in Zuvekas and Cohen (2002), and also can be accessed via the CFACT Data Center. For more information, see the Data Center section of the MEPS website. If examining trends in MEPS expenditures, please refer to Section 3.5 for more information.

Data Editing and Imputation Methodologies of Expenditure Variables

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

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

General Data Editing Methodology

Logical edits were used to resolve internal inconsistencies and other problems in the HC and the MPC 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.

Imputation Methodologies

The predictive mean matching imputation method was used to impute missing expenditures. This procedure uses regression models (based on events with completely reported expenditure data) to predict total expenses for each event. Then, for each event with missing payment information, a donor event with the closest predicted payment with the same pattern of expected payment sources as the event with the missing payment was used to impute the missing payment value.

A weighted sequential hot-deck procedure was used to impute the missing total charges. This procedure uses survey data from respondents to replace missing data while using weights to ensure that the results are representative of the entire population.

Home Health Data Editing and Imputation

Expenditures for home health events were developed in a sequence of logical edits and imputations. (Analysts should note that home health care provided by friends, family, or volunteers was assumed not to have associated expenditures and was not included in any imputation process. All expenditures for home health care provided by informal care providers were assigned Inapplicable (-1) because those types of events were skipped out of (never asked) the questions regarding expenditures. Household edits were applied to sources and amounts of

payment for all household-reported events for paid independent providers and unmatched agency providers. MPC edits were applied to provider-reported sources and amounts of payment for records matched to household-reported events for all agency home health providers. Both sets of edits were used to correct obvious errors in the reporting of expenditures. Imputations for independent paid providers and for agencies were conducted separately. Logical edits were used to sort each event into a specific category for the imputations. Events with complete expenditures were flagged as potential donors while events with missing expenditure data were assigned to various recipient categories based on the extent of their 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 separate recipient categories.

Expenditures were imputed using a predictive mean matching method. The donor pool in these imputations includes events with complete expenditures from the HC for paid independent providers (HHP) and restricted to the MPC for agency providers (HHA). As stated previously, home health care provided by friends, family, or volunteers (informal, MPCELIG = 3) was assumed not to have expenditures associated with it and was not included in any imputation process.

Imputation Flag Variable (IMPFLAG)

IMPFLAG is a six-category variable that indicates if the event contains complete HC or 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 events)

IMPFLAG = 1 complete HC data

IMPFLAG = 2 complete MPC data

IMPFLAG = 3 fully imputed

IMPFLAG = 4 partially imputed

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

Flat Fee Expenditures

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. Because MEPS does not collect flat fee information about home health events, there are no flat fee expenditure data included in this PUF.

Zero Expenditures

There are some medical events reported by respondents for which the payments were zero. This could occur for several reasons including (1) free care was provided, (2) bad debt was incurred, (3) follow-up events were provided without a separate charge (e.g., after a surgical procedure), or (4) the event was paid for through government or privately-funded research or clinical trials. If all of the medical events for a person fell into one of these categories, then the total annual expenditures for that person would be zero. All expenditures for home health care provided by informal care providers (family, friends, or volunteers, MPCELIG = 3) were assigned Inapplicable (-1) because those types of events were skipped out of (never asked) questions regarding expenditures.

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) - includes any deductible, coinsurance, and copayment amounts not covered by other sources, as well as payments for services and providers not covered by the person's insurance or other sources,
2. Medicare,
3. Medicaid,
4. Private insurance,
5. Veterans administration/CHAMPVA, 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.

Home Health Expenditure Variables (HHSF22X - HHXP22X)

Home health agency, hospital, and nursing home events are sampled at a rate of 100% for the MPC. Households were not asked any expenditure-related questions regarding these types of events; therefore, there are no household-reported expenditure data for these events. Conversely,

paid independent providers are not included in the MPC. Household-reported responses are the only data available for these types of events. All expenditure data for paid independent providers are fully imputed from household-reported expenditures. There are no expenditure data for informal care providers. Informal care (MPCELIG = 3, unpaid care provided by family, friends, or volunteers) was assigned Inapplicable (-1), in all expenditure categories.

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

All of these expenditures have gone through an editing and imputation process and have been rounded to the nearest penny. HHSF22X - HHOT22X are the 10 sources of payment. HHXP22X is the sum of the 10 sources of payment for the home health expenditures, and HHTC22X is the total charge. The 10 sources of payment are: self/family (HHSF22X), Medicare (HHMR22X), Medicaid (HHMD22X), private insurance (HHPV22X), Veterans Administration/CHAMPVA (HHVA22X), TRICARE (HHTR22X), other federal sources (HHOF22X), state and local (non-federal) government sources (HHSL22X), Workers' Compensation (HHWC22X), and other insurance (HHOT22X). Analysts can determine if a home health event was provided by an agency or by some other paid independent provider by subsetting the variable MPCELIG to the appropriate and desired value.

Rounding

Expenditure variables on the 2022 HHV PUF have been rounded to the nearest penny. Person-level expenditure information to be released on the 2022 Consolidated PUF will be rounded to the nearest dollar. It should be noted that using the 2022 MEPS event files to create person-level totals will yield slightly different totals than those on the Consolidated PUF. These differences are due to rounding only. Moreover, in some instances, the number of persons with expenditures in the event PUFs for a particular source of payment may differ from the number of persons with expenditures on the person-level expenditure PUF for that source of payment. This difference is also an artifact of rounding only.

3.0 Survey Sample Information

3.1 Discussion of Pandemic Effects on Quality of MEPS Data

The challenges associated with MEPS data collection in 2020 after the onset of the COVID-19 pandemic continued through 2021 and possibly into 2022. The major modifications to the standard MEPS study design remained in effect, permitting data to be collected safely but with accompanying concerns related to the quality of the data obtained. The suggestion made in the documentation for the FY 2020 and FY2021 MEPS Consolidated PUF data still holds. Researchers are counseled to take care in the interpretation of estimates based on data collected from these three calendar years. This includes the comparison of such estimates to those of other years and corresponding trend analyses.

Section 3.1 of the documentation for the 2020 Consolidated PUF provides a general discussion of the impact of the COVID-19 pandemic on several other major in-person federal surveys as well as on MEPS. In addition, it offers a detailed look at how MEPS was modified to permit safe data collection and the development of useful estimates at a time when the way the U.S. health care system functioned underwent many transformations to meet population needs. Three sources of potential bias were identified for MEPS for FY 2020: (1) long recall period for Round 6 of Panel 23, (2) switching from in-person to telephone interviewing which likely had a larger impact on Panel 25, and (3) the impact of CPS bias on the MEPS weights. A number of statistically significant differences were found between panels for FY 2020. Those findings are discussed in MEPS HC-224.

Concerns of potential bias for FY 2021 and between panel differences are discussed in Section 3.1 of the documentation for the 2021 Consolidated PUF. Additional analysis has also uncovered a concerning trend on event reporting in MEPS following the COVID-19 pandemic. While reporting of other event types has rebounded from the dip experienced in 2020, inpatient (IP) and emergency room (ER) utilization reports collected in FY 2021 did not rebound as much as key benchmarks, even though these are the most salient event types. Modifications made to the MEPS sample design discussed in the 2022 Population Characteristics PUF may have partially contributed to the concerning trend.

Concerns for potential bias for FY 2022 include:

- The impact of the pandemic on NHIS data collection and the resulting Panel 26 MEPS sample (Section 3.1.1 of the 2022 Population Characteristics PUF). NHIS response rates in the pandemic and shifts in the resulting MEPS sample may have increased the likelihood that the MEPS Panel 26 respondents differed in composition compared to previous years.

- The extension of panels (beginning of Section 3.1 of the 2022 Population Characteristics PUF). While there is a benefit in boosting the MEPS sample size by keeping pre-pandemic panels active for an additional two years to counter reduced response rates, there are two risks with this approach: attrition in these panels beyond what is experienced in two years, which may lead households with more serious health issues to leave MEPS, and a conditioning effect whereby respondents learn over time that reporting events results in a longer interview.

- Significantly lower response rates (Section 3.2 of the 2022 Population Characteristics PUF) that could differentially exclude households more likely to experience IP stays. The demographic shifts on MEPS between 2019 and 2021 suggest a more educated, higher-income, older MEPS.

Preliminary analyses undertaken to examine the quality of the MEPS FY 2022 data compared health care utilization for the MEPS target population between the panels fielded. These comparisons were undertaken for the full sample and the three age groups of 0-17, 18-64, and 65+.

These comparisons found no major differences in IP or ER visits between the three panels. Slight differences were observed in dental visits and outpatient visits. For dental visits, Panel 26 reported at a higher rate than Panel 24 or Panel 27 in the age range 18-64. For outpatient visits, Panel 24 reported at a lower rate than Panel 26 and Panel 27 in the age range 18-64.

In summary, the weights developed for the MEPS FY 2022 data can be expected to produce useful estimates for initial analyses. Further analyses of MEPS estimates will be conducted as part of the production of the FY 2022 Consolidated PUF to be released later in 2024. This will help identify any additional data quality issues as well as possible improvements that could be implemented.

The various actions taken in the development of the person-level weights for the MEPS FY 2022 data were designed to limit the potential for bias in the data due to changes in data collection and response bias. However, evaluations of MEPS data quality in 2021 and 2022 suggest that users of the MEPS FY 2022 PUFs should continue to exercise caution when interpreting estimates and assessing analyses based on these data, as well as in comparing 2022 estimates to those of prior years.

3.2 Sample Weight (PERWT22F)

There is a single full-year person-level weight (PERWT22F) assigned to each record for each Key, in-scope person who responded to MEPS for the full period of time that they were in scope during 2022. A Key person was either a member of a responding NHIS household at the time of the interview or joined 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 those returning from military service, an institution, or residence in a foreign country). A person is in scope whenever they are a member of the civilian noninstitutionalized portion of the U.S. population.

3.3 Details on Person Weight Construction

The person-level weight PERWT22F was developed in several stages. First, a person-level weight for Panel 24 was created, including an adjustment for nonresponse over time and raking. The raking involved adjusting to several sets of marginal control totals reflecting Current Population Survey (CPS) population estimates based on six variables. The six variables used in the establishment of the initial person-level control figures were: educational attainment of the reference person (three categories: no degree; high school/GED only or some college; bachelor's or a higher degree); Census region (Northeast, Midwest, South, West); MSA status (MSA, non-MSA); race/ethnicity (Hispanic; Black, non-Hispanic; Asian, non-Hispanic; and other); sex; and age (0-18, 19-25, 26-34, 35-44, 45-64, and 65 or older). (Note, however, that for confidentiality reasons, the MSA status variables are no longer released for public use.) The person-level weights for Panel 26 and Panel 27 were created similarly. Secondly, a composite weight was formed by multiplying each weight from Panel 24 by the factor .22, each weight from Panel 26 by the factor .29, and each weight from Panel 27 by the factor .49. The choice of factors reflected the relative effective sample sizes of the three panels, helping to limit the variance of estimates

obtained from pooling the three samples. Weights for the 2022 Population Characteristics PUF were then developed by raking the composite weight to the same set of CPS-based control totals.

The approach for establishing the 2022 Consolidated PUF weight is as follows. When poverty status information derived from MEPS income variables becomes available, a final raking is undertaken. The full sample weight appearing on the Population Characteristics PUF for a given year is re-raked, replacing educational attainment with poverty status while retaining the other five raking variables previously indicated. Specifically, control totals based on CPS estimates of 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 age, race/ethnicity, sex, region, and MSA status are used to calibrate weights.

3.3.1 MEPS Panel 24 Weight Development Process

The person-level weight for MEPS Panel 24 was developed using the 2021 full-year weight for an individual as a "base" weight for 2021 survey participants present in 2022. For Key, in-scope members who joined an RU some time in 2022 after being out of scope in 2021, the initially assigned person-level weight was the corresponding 2021 family weight. The weighting process included an adjustment for person-level nonresponse over Rounds 8 and 9 as well as raking to population control figures for December 2022 for Key, responding persons in scope on December 31, 2022. These control totals were derived by scaling back the population distribution obtained from the March 2023 CPS to reflect the December 31, 2022 estimated population total (estimated based on Census projections for January 1, 2023). Variables used for person-level raking included: education of the reference person (three categories: no degree; high school/GED only or some college; bachelor's or a higher degree); Census region (Northeast, Midwest, South, West); MSA status (MSA, non-MSA); race/ethnicity (Hispanic; Black, non-Hispanic; Asian, non-Hispanic; and other); sex; and age (0-18, 19-25, 26-34, 35-44, 45-64, and 65 or older). (Note, however, that for confidentiality reasons, the MSA status variables are no longer released for public use.) The final weight for Key, responding persons who were not in scope on December 31, 2022 but were in scope earlier in the year was the nonresponse-adjusted person weight without raking.

The 2021 full-year weight used as the base weight for Panel 24 was derived from the 2019 MEPS Round 1 weight and reflected adjustment for nonresponse over the remaining data collection rounds in 2019, 2020, and 2021 as well as raking to the December 2019, December 2020, and December 2021 population control figures.

3.3.2 MEPS Panel 26 Weight Development Process

The person-level weight for MEPS Panel 26 was developed by using the 2021 full-year weight as a "base" weight for survey participants present in 2022.

For Key, in-scope members who joined an RU at some time in 2022 after being out of scope in 2021, the initially assigned person-level weight was the corresponding 2021 family weight. The weighting process also included an adjustment for person-level nonresponse over Rounds 4 and 5 as well as raking to the same population control figures for December 2022 used for the Panel

24 weight for Key, responding persons in scope on December 31, 2022. The same six variables used for Panel 24 raking (education level, Census region, MSA status, race/ethnicity, sex, and age) were also used for Panel 26 raking. Similar to Panel 24, the Panel 26 final weight for Key, responding persons not in scope on December 31, 2022 but in scope earlier in the year was the nonresponse-adjusted person weight without raking.

Note that the 2021 full-year weight that was used as the base weight for Panel 26 was derived using the 2021 MEPS Round 1 weight and reflected adjustment for nonresponse over the remaining data collection rounds in 2021 as well as raking to the December 2021 population control figures.

3.3.3 MEPS Panel 27 Weight Development Process

The person-level weight for Panel 27 was developed using the 2022 Round 1 person-level weight as a "base" weight. The Round 1 weights incorporated the following components: the original household probability of selection for the NHIS and for the NHIS subsample reserved for the MEPS, an adjustment for NHIS nonresponse, the probability of selection for MEPS from the NHIS, an adjustment for nonresponse at the dwelling unit level for Round 1, and raking to control figures at the person level obtained from the March CPS of the corresponding year. For Key, in-scope members who joined an RU after Round 1, the Round 1 DU weight served as a "base" weight.

The weighting process also included an adjustment for nonresponse over the remaining data collection rounds in 2022 as well as raking to the same population control figures for December 2022 that were used for the Panel 24 and Panel 26 weights for Key, responding persons in scope on December 31, 2022. The same six variables used for Panel 24 and Panel 26 raking (education level of the reference person, Census region, MSA status, race/ethnicity, sex, and age) were also used for Panel 27 raking. Similar to Panel 24 and Panel 26, the Panel 27 final weight for Key, responding persons who were not in scope on December 31, 2022 but were in scope earlier in the year was the nonresponse-adjusted person weight without raking.

3.3.4 The Final Weight for 2022

The final raking of those in scope at the end of the year has been described above. In addition, the composite weights of two groups of persons who were out of scope on December 31, 2022 were adjusted for expected undercoverage. Specifically, the weights of those who were out of scope on December 31, 2022, but in scope at some time during the year and were residing in a nursing home at the end of the year were poststratified to an estimate of the number of persons who were residents of Medicare- and Medicaid-certified nursing homes for part of the year (approximately 3-9 months) during 2014. This estimate was developed from data on the Minimum Data Set (MDS) of the Center for Medicare and Medicaid Services (CMS). The weights of persons who died while in scope were poststratified to corresponding estimates derived using data obtained from the Centers for Disease Control and Prevention (CDC), National Center for Health Statistics (NCHS), Provisional Mortality Statistics, 2018 through Last Week on CDC WONDER Online Database, released in 2023, the latest available data at the

time. Separate decedent control totals were developed for the "65 and older" and "under 65" civilian noninstitutionalized populations.

Overall, the weighted population estimate for the civilian noninstitutionalized population for December 31, 2022 is 329,059,733 (PERWT22F >0 and INSC1231=1). The sum of person-level weights across all persons assigned a positive person-level weight is 333,053,243.

3.4 Coverage

The target population associated with MEPS is the 2022 U.S. civilian noninstitutionalized population. However, the MEPS sampled households are a subsample of the NHIS households interviewed in 2018 (Panel 24), 2020 (Panel 26), and 2021 (Panel 27). New households created after the NHIS interviews for the respective panels and consisting exclusively of persons who entered the target population after 2018 (Panel 24), after 2020 (Panel 26), or after 2021 (Panel 27) are not covered by the 2022 MEPS. Nor are previously out of scope persons who joined an existing household but are not related 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. Those not covered represent a small proportion of the MEPS target population.

3.5 Using MEPS Data for Trend Analysis

For analysts using the MEPS data for trend analysis, we note that there are uncertainties associated with 2020, 2021, and possibly 2022 data quality for reasons discussed throughout Section 3. Preliminary evaluations of a set of MEPS estimates of particular importance suggest that they are of reasonable quality. Nevertheless, analysts are advised to exercise caution in interpreting these estimates, particularly in terms of trend analyses, since access to health care was substantially affected by the COVID-19 pandemic, as were related factors such as health insurance and employment status for many persons.

The MEPS began in 1996, and the utility of the survey for analyzing health care trends expands with each additional year of data; however, when examining trends over time using the MEPS, the length of time being analyzed should 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, economic conditions, or the MEPS methodology.

With respect to methodological considerations, changes in data collection methods, such as interviewer training, were introduced in 2013 to obtain more complete information about health care utilization from MEPS respondents; the changes were fully implemented in 2014. This effort likely resulted in improved data quality and a reduction in underreporting starting in the second half of 2013 and continuing throughout 2014 full year files; the changes have also had some impact on analyses involving trends in utilization across years. The changes in the NHIS sample design in 2016 and 2018 could also potentially affect trend analyses. The new NHIS

sample design is based on more up-to-date information related to the distribution of housing units across the United States. As a result, it can be expected to better cover the full civilian noninstitutionalized population, the target population for MEPS, as well as many of its subpopulations. Better coverage of the target population helps to reduce the potential for bias in both NHIS and MEPS estimates.

Another change with the potential to affect trend analyses involved major modifications to the MEPS instrument design and data collection process, particularly in the events sections of the instrument. These were introduced in the spring of 2018 and thus affected data beginning with Round 1 of Panel 23, Round 3 of Panel 22, and Round 5 of Panel 21. Since the full year 2017 MEPS files were established from data collected in Rounds 1-3 of Panel 22 and Rounds 3-5 of Panel 21, they reflected two instrument designs. To mitigate the effect of such differences within the same full-year file, the Panel 22 Round 3 data and the Panel 21 Round 5 data were transformed to make them as consistent as possible with data collected under the previous design. The changes in the instrument were designed to make the data collection effort more efficient and easier to administer. In addition, expectations were that data on some items, such as those related to health care events, would be more complete with the potential of identifying more events. Increases in service use reported since the implementation of these changes are consistent with these expectations. *Analysts should be aware of the possible impacts of these changes on the data and especially on trend analyses that include the year 2018 because of the design transition.*

Process changes, such as data editing and imputation, may also affect trend analyses. For example, users should refer to Section 2.5.11: Utilization, Expenditures, and Sources of Payment Variables in the Consolidated PUF (HC-243) and, for more detail, to the documentation for the prescription drug file (HC-239A) when analyzing prescription drug spending over time.

As always, it is recommended that, before conducting trend analyses, analysts should review relevant sections of the documentation for descriptions of these types of changes that might affect the interpretation of changes over time.

To smooth or stabilize trend analyses based on the MEPS data, analysts may also wish to consider using statistical techniques such as comparing pooled time periods (e.g. 1996-1997 versus 2011-2012), working with moving averages, or using modeling techniques with several consecutive years of the data.

Finally, statistical significance tests should be conducted to assess the likelihood that observed trends are not attributable to sampling variation. In addition, researchers should be aware of the impact of multiple comparisons on Type I error. Without making appropriate allowance for multiple comparisons, the use of numerous statistical significance tests of trends will increase the likelihood of concluding that a change has taken place when one has not.

4.0 Strategies for Estimation

4.1 Developing Event-Level Estimates

The data in this PUF can be used to develop national 2022 event-level (i.e., monthly) estimates for the U.S. civilian noninstitutionalized population on expenditures and sources of payment for home health care medical provider visits. The weight assigned to each home health care medical provider event reported is the person-level weight of the person who was visited. If a person had several events reported, each event is assigned that individual's person-level weight. Estimates must be weighted by PERWT22F to be nationally representative. For example, the appropriate estimate for the overall mean out-of-pocket payment per month of care is computed as follows (the subscript 'j' identifies each event and represents a numbering of events from 1 through the total number of events in the file):

$(\sum W_j X_j)/(\sum W_j)$, where

$W_j = PERWT22F_j$ (full-year person weight for the person associated with event j) and

$X_j = HHSF22X_j$ (amount paid by self/family for event j)

Estimates and corresponding standard errors (SE) can be derived using an appropriate computer software package for complex survey analysis such as SAS, Stata, SUDAAN, R, or SPSS.

The tables below contain the event-level estimates for several key variables on this file. Informal care (MPCELIG = 3) is not included in the tables because, by definition, there are no payments for those events and, therefore, no expenditure data are collected.

Table 3

Selected Event-Level Estimates - Expenditures: Home Health Agency & Paid Independents (MPCELIG = 1, 2)

Estimate of Interest	Variable	Estimate (SE)	Estimate Excluding Zero Payment Events (SE)
Proportion of events with expenditures > 0*	HHXP22X	0.983 (0.0045)	_____
Mean total payments per month of care	HHXP22X	$1,663 (183.9000)	$1,692 (186.6000)
Mean out-of-pocket payments per month of care	HHSF22X	$99 (23.3000)	$101 (23.7000)
Mean proportion of total monthly expenditures paid out of pocket	HHSF22X/ HHXP22X	_____	0.127 (0.0184)

Table 4

Selected Event-Level Estimates - Expenditures: Home Health Agency Providers only (MPCELIG = 1)

Estimate of Interest	Variable	Estimate (SE)	Estimate Excluding Zero Payment Events (SE)
Proportion of events with expenditures > 0*	HHXP22X	0.981 (0.0053)	_____
Mean total payments per month of care	HHXP22X	$1,740 (214.3000)	$1,774 (217.9000)
Mean out-of-pocket payments per month of care	HHSF22X	$15 (6.3000)	$16 (6.4000)
Mean proportion of total monthly expenditures paid out of pocket	HHSF22X/ HHXP22X	_____	0.027 (0.0090)

Table 5

Selected Event-Level Estimates - Expenditures: Paid Independent Providers only (MPCELIG = 2)

Estimate of Interest	Variable	Estimate (SE)	Estimate Excluding Zero Payment Events (SE)
Proportion of events with expenditures > 0*	HHXP22X	0.996 (0.0024)	_____
Mean total payments per month of care	HHXP22X	$1,252 (202.6000)	$1,256 (203.5000)
Mean out-of-pocket payments per month of care	HHSF22X	$548 (121.3000)	$550 (121.8000)
Mean proportion of total monthly expenditures paid out of pocket	HHSF22X/ HHXP22X	_____	0.655 (0.0632)

*Zero payment events can occur in MEPS for the following reasons: (1) there was no charge for a follow-up event, (2) the provider was never paid by an individual, insurance plan, or other source for services provided, (3) the charges were included in another bill, or (4) the event was paid for through government or privately-funded research or clinical trials.

4.2 Person-Based Estimates for Home Health Care

To enhance analyses of home health care, analysts may link information about the home health care received by sample persons in this file to the Consolidated PUF (which has data for all MEPS sample persons), or conversely, link person-level information from the Consolidated PUF to this event-level file. Both this file and the Consolidated PUF may be used to derive estimates

relative to persons with home health care and annual estimates of total expenditures. However, for estimates that pertain to those who did not receive home health care as well as those who did (for example, the percentage of adults with at least one month in which home health care was provided during the past year or the mean number of home health care visits in the past year among those 65 or older), this file cannot be used. Only those persons with at least one month in which home health care was provided are represented on this data file. The Consolidated PUF must be used for person-level analyses that include both those with and without home health care.

4.3 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 negative 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 software package 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 and zero expenditures) are described in "Data Editing and Imputation Methodologies of Expenditure Variables."

4.4 Variance Estimation (VARPSU, VARSTR)

To obtain estimates of variability in the MEPS estimates (such as the standard error of sample estimates or corresponding confidence intervals) analysts should take into account the complex sample design of the MEPS for both person-level and family-level analyses. Several methodologies have been developed for estimating standard errors for surveys with a complex sample design, including the Taylor-series linearization method, balanced repeated replication (BRR), and jackknife replication. Various software packages provide analysts with the capability of implementing these methodologies. MEPS analysts most commonly use the Taylor-series approach. Although this PUF does not contain replicate weights, analysts can use the BRR methodology to construct replicate weights to develop variances for more complex estimators (see Section 4.4.2).

4.4.1 Taylor-series Linearization Method

The variables needed to calculate appropriate standard errors based on the Taylor-series linearization method are included on this file and all other MEPS PUFs. Software packages that permit the use of the Taylor-series linearization method include SUDAAN, Stata, R, SAS (version 8.2 and higher), and SPSS (version 12.0 and higher). For complete information on the capabilities of a package, analysts should refer to the user documentation for the software.

With the Taylor-series linearization method, variance estimation strata and the variance estimation PSUs within these strata must be specified. The variables VARSTR and VARPSU on this HHV PUF identify the sampling strata and primary sampling units required by the variance estimation programs. Specifying a "with replacement" design in one of the previously mentioned software packages will provide estimated standard errors appropriate for assessing the variability of the MEPS 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 number available. For variables of interest distributed throughout the country (and thus the MEPS sample PSUs), one can generally expect to see at least 100 degrees of freedom associated with the estimated standard errors for national estimates based on this MEPS database.

Before 2002, the 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. Beginning with the 2002 point-in-time PUF, the approach changed with the intention that variance strata and PSUs would be developed to be compatible with all future PUFs until the NHIS design changed. Thus, when pooling data across years 2002 through Panel 11 of the 2007 files, analysts can use the variance strata and PSU variables provided without modifying them 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.

Beginning in Panel 12 of the 2007 files, a new set of variance strata and PSUs was developed because of the introduction of a new NHIS design. There are 165 variance strata with either two or three variance estimation PSUs per stratum. Therefore, there are a total of 368 (203+165) variance strata in the 2007 Population Characteristics PUF, as it consists of two panels that were selected under two independent NHIS sample designs. Since both MEPS panels in the full-year files from 2008 through 2016 were based on the same NHIS design, there were only 165 variance strata. These variance strata (VARSTR values) have been numbered from 1001 to 1165 so that they can be readily distinguished from those developed under the former NHIS sample design if data are pooled for several years.

The NHIS sample design was changed again in 2016, effectively changing the MEPS design beginning with calendar year 2017. Beginning in Panel 22 of the 2017 files, a new set of variance strata and PSUs were developed. There are 117 variance strata with either two or three variance estimation PSUs per stratum. Therefore, there are a total of 282 (165+117) variance strata in the 2017 Population Characteristics PUF, as it consisted of two panels that were selected under two independent NHIS sample designs. To make the pooling of data across multiple years of the MEPS more straightforward, the numbering system for the variance strata was changed. The strata associated with the new design are numbered from 2001 to 2117.

The new NHIS sample design was further modified in 2018, so the MEPS variance structure for the 2019 Population Characteristics PUF was also modified, reducing the number of variance strata to 105. Consistency was maintained with the prior structure in that the 2019 variance strata were also numbered within the range of values from 2001 to 2117, although there are now gaps in the values assigned within this range. Because of the modification, each stratum could contain up to 5 variance estimation PSUs.

For Panel 26 in the 2021 and 2022 Population Characteristics PUF, an additional NHIS sample was used for the MEPS to account for increasing nonresponse during the pandemic (as discussed in Section 3.1). The additional sample was assigned to the existing variance strata, so the Population Characteristics PUF continues to have 105 variance strata, numbered 2001-2117, with a few gaps in the values in that range. In many cases, the additional sample was assigned to new variance estimation PSUs; thus, in the Population Characteristics PUF, each stratum contains up to eight variance estimation PSUs.

Some analysts may be interested in pooling data across multiple years of MEPS data. When doing so, analysts should note that to obtain appropriate standard errors it is necessary to specify a common variance structure. Before 2002, each annual PUF was released with a variance structure unique to the particular MEPS sample in that year. Starting in 2002, the annual PUFs were released with a common variance structure that allowed users to pool data from 2002 through 2018. However, analysts can no longer do this routinely because the variance structure had to be modified beginning with 2019.

To ensure that variance strata are identified appropriately for variance estimation purposes when pooling MEPS data across several years, analysts can proceed as follows:

1. When pooling any year from 2002 through 2018, use the variance strata numbering as is.

2. When pooling (a) any year from 1996 to 2001 with any year from 2002 or later, or (b) the year 2019 and beyond with any earlier year, use the pooled linkage PUF HC-036, which contains the proper variance structure. The HC-036 file is updated every year so that appropriate variance structures are available with pooled data. Further details on the HC-036 file are included in the public use documentation of the HC-036 file.

4.4.2 Balanced Repeated Replication Method

BRR replicate weights are not provided on this MEPS PUF for the purposes of variance estimation. However, a file containing a BRR replication structure is made available so that the users can form replicate weights, if desired, from the final MEPS weight to compute variances of MEPS estimates using either BRR or Fay's modified BRR (Fay, 1989) methods. The replicate weights are useful for computing variances of complex non-linear estimators for which a Taylor linear form is neither easy to derive nor available in commonly used software. For instance, it is not possible to calculate the variances of a median or the ratio of two medians by using the Taylor linearization method. For these types of estimators, users can calculate a variance using BRR or Fay's modified BRR methods. However, it should be noted that the replicate weights have been derived from the final weight through a shortcut approach. Specifically, the replicate weights are not computed starting with the base weight, and all adjustments made in different stages of weighting are not applied independently in each replicate. Thus, the variances computed by using this one-step BRR do not capture the effects of all weighting adjustments that would be captured in a set of fully developed BRR replicate weights. The Taylor series approach does not fully capture the effects of the different weighting adjustments either.

The dataset, HC-036BRR, MEPS 1996-2021 Replicates for Variance Estimation File, contains the information necessary to construct the BRR replicates. It includes a set of 128 flags (BRR1-BRR128) in the form of half sample indicators, each of which is coded 0 or 1 to indicate whether the person should or should not be included in that particular replicate. These flags can be used in conjunction with the full-year weight to construct the BRR replicate weights. For an analysis of MEPS data pooled across years, the BRR replicates can be formed in the same way by using the HC-036, MEPS 1996-2021 Pooled Linkage Variance Estimation File. For more information about creating BRR replicates, users can refer to the documentation for the HC-036BRR pooled linkage file on the AHRQ website.

5.0 Merging/Linking MEPS Data Files

Data from this PUF can be used alone or in conjunction with other PUFs for different analytic purposes. Merging characteristics of interest from other MEPS PUFs expands the scope of potential estimates. For example, the medical event PUFs can be merged with the person-level Consolidated PUF to calculate event-level estimates for persons with specific characteristics (e.g., age, race, sex, and education).

Most of the event PUFs can also be linked to the Medical Conditions PUF by using the condition-event link (CLNK) PUF. When using the CLNK PUF, analysts should keep in mind that (1) conditions are household reported, (2) there may be multiple conditions associated with a medical event, (3) one condition may link to more than one event, and (4) not all medical events link to the Medical Conditions PUF.

In addition to linking to other MEPS PUFs, each MEPS panel can also be linked back to the previous year's NHIS public use files. This is because the set of households selected for MEPS is a subsample of those participating in the NHIS. For information on obtaining MEPS/NHIS link files please see the MEPS website.

References

Bramlett, M.D., Dahlhamer, J.M., & Bose, J. (2021, September). *Weighting Procedures and Bias Assessment for the 2020 National Health Interview Survey*. Centers for Disease Control and Prevention.

Chowdhury, S.R., Machlin, S.R., & Gwet, K.L. Sample designs of the Medical Expenditure Panel Survey Household Component, 1996-2006 and 2007-2016. (2019, January) Methodology Report #33. Rockville, MD: Agency for Healthcare Research and Quality.

Cohen, S.B. (1996). The redesign of the Medical Expenditure Panel Survey: A component of the DHHS survey integration plan. *Proceedings of the Council of Professional Associations on Federal Statistics Seminar on Statistical Methodology in the Public Service.*

Cox, B.G. and Cohen, S.B. (1985). *Methodological issues for health care surveys.* Marcel Dekker.

Dahlhamer, J.M., Bramlett, M.D., Maitland, A., & Blumberg, S.J. (2021). Preliminary evaluation of nonresponse bias due to the COVID-19 pandemic on National Health Interview Survey estimates, April-June 2020. Hyattsville, MD: National Center for Health Statistics.

Fay, R.E. (1989). Theory and application of replicate weighting for variance calculations. *Proceedings of the Survey Research Methods Sections of the American Statistical Association*, 212-217.

Lau, D.T., Sosa, P., Dasgupta, N., & He, H. (2021). Impact of the COVID-19 pandemic on public health surveillance and survey data collections in the United States. *American Journal of Public Health, 111* (12), 2118-2121.

Monheit, A.C., Wilson, R., and Arnett, III, R.H. (Eds.). (1999). *Informing American health care policy*. Jossey-Bass Inc.

Rothbaum, J. & Bee, A. (2021, May 3). *Coronavirus infects surveys, Too: Survey nonresponse bias and the coronavirus pandemic*. Washington, DC: U.S. Census Bureau.

Rothbaum, J. & Bee, A. (2022, September 13). *How has the pandemic continued to affect survey response? Using administrative data to evaluate nonresponse in the 2022 Current Population Survey Annual Social and Economic Supplement*. Washington, DC: U.S. Census Bureau.

RTI International (2019). *Medical Provider Component (MEPS-MPC) methodology report 2017 data collection.* Rockville, MD. Agency for Healthcare Research and Quality.

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 Institute.

U.S. Census Bureau. *Current Population Survey: 2021 Annual Social and Economic (ASEC) Supplement*. (2021). Washington, DC: Author.

Zuvekas, S.H. & Cohen, J.W. (2002). A guide to comparing health care expenditures in the 1996 MEPS to the 1987 NMES. *Inquiry*. 39(1), 76-86.

Zuvekas, S.H. & Kashihara, D. (2021). The impacts of the COVID-19 pandemic on the Medical Expenditure Panel Survey. *American Journal of Public Health, 111 (12)*, 2157-2166.

D. Variable-Source Crosswalk

FOR MEPS HC 239H: 2022 HOME HEALTH VISITS

Survey Administration Variables

Variable	Description	Source
DUID	Panel # + Encrypted DU identifier	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
PANEL	Panel Number	Constructed

Home Health Events Variables

Variable	Description	Source
HHDATEYR	Event date - year	CAPI derived
HHDATEMM	Event date - month	CAPI derived
MPCELIG	MPC eligibility flag	Constructed
SELFAGEN	Does provider work for agency or self	EV60
HHTYPE	Home health event type	EV50
CNA_M18	Type of prof hlth care wrkr - cert nurse asst	HH10
DIETICN_M18	Type of prof hlth care wrkr - dietitian/nutrt	HH10
IVTHP_M18	Type of prof hlth care wrkr - iv or infusion therapist	HH10
MEDLDOC_M18	Type of prof hlth care wrkr - medical doctor	HH10
NURPRACT_M18	Type of prof hlth care wrkr - nurse/practr	HH10
OCCUPTHP_M18	Type of prof hlth care wrkr - occupational therap	HH10

Variable	Description	Source
PHYSLTHP_M18	Type of prof hlth care wrkr - physical therapy	HH10
RESPTHP_M18	Type of prof hlth care wrkr - respira therapy	HH10
SOCIALW_M18	Type of prof hlth care wrkr - social worker	HH10
SPEECTHP_M18	Type of prof hlth care wrkr - speech therapy	HH10
HCarWrkrProfNone_M18	None of the listed professional home health providers	HH10
COMPANN_M18	Type of non prof hlth care wrkr - companion	HH20
HMEMAKER_M18	Type of non prof hlth care wrkr - homemaker/house cleaner	HH20
HHAIDE_M18	Type of non prof hlth care wrkr - home health / care aide	HH20
HOSPICE_M18	Type of non prof hlth care wrkr - hospice worker	HH20
NURAIDE_M18	Type of non prof hlth care wrkr - nurse's aide	HH20
PERSONAL_M18	Type of non prof hlth care wrkr - pers care attdt	HH20
HCarWrkrNonProfNone_M18	None of the listed non professional home health providers	HH20
VSTRELCN	Any hh care svce related to hlth cond	HH70
FREQCY	Provider helped every week/some weeks	HH90
DAYSPWK	# days / week provider came	HH100
DAYSPMO	# days / month provider came	HH110
SAMESVCE_M18	Any oth mons per received same services	HH120
HHDAYS	Days per month in home health, 2022	Constructed

Imputed Expenditure Variables

Variable	Description	Source
HHSF22X	Amount paid, family (Imputed)	CP Section (Edited)
HHMR22X	Amount paid, Medicare (Imputed)	CP Section (Edited)
HHMD22X	Amount paid, Medicaid (Imputed)	CP Section (Edited)
HHPV22X	Amount paid, private insurance (Imputed)	CP Section (Edited)
HHVA22X	Amount paid, Veterans/CHAMPVA (Imputed)	CP Section (Edited)
HHTR22X	Amount paid, TRICARE (Imputed)	CP Section (Edited)
HHOF22X	Amount paid, other federal (Imputed)	CP Section (Edited)
HHSL22X	Amount paid, state & local gov (Imputed)	CP Section (Edited)
HHWC22X	Amount paid, workers comp (Imputed)	CP Section (Edited)
HHOT22X	Amount paid, other insurance (Imputed)	CP Section (Edited)
HHXP22X	Sum of HHSF22X - HHOT22X (Imputed)	Constructed
HHTC22X	Household reported total charge (Imputed)	CP Section (Edited)
IMPFLAG	Imputation status	Constructed

Weights Variables

Variable	Description	Source
PERWT22F	Expenditure file person weight, 2022	Constructed
VARSTR	Variance estimation stratum, 2022	Constructed
VARPSU	Variance estimation PSU, 2022	Constructed