MEPS HC 239F:
2022 Outpatient Department Visits

June 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.

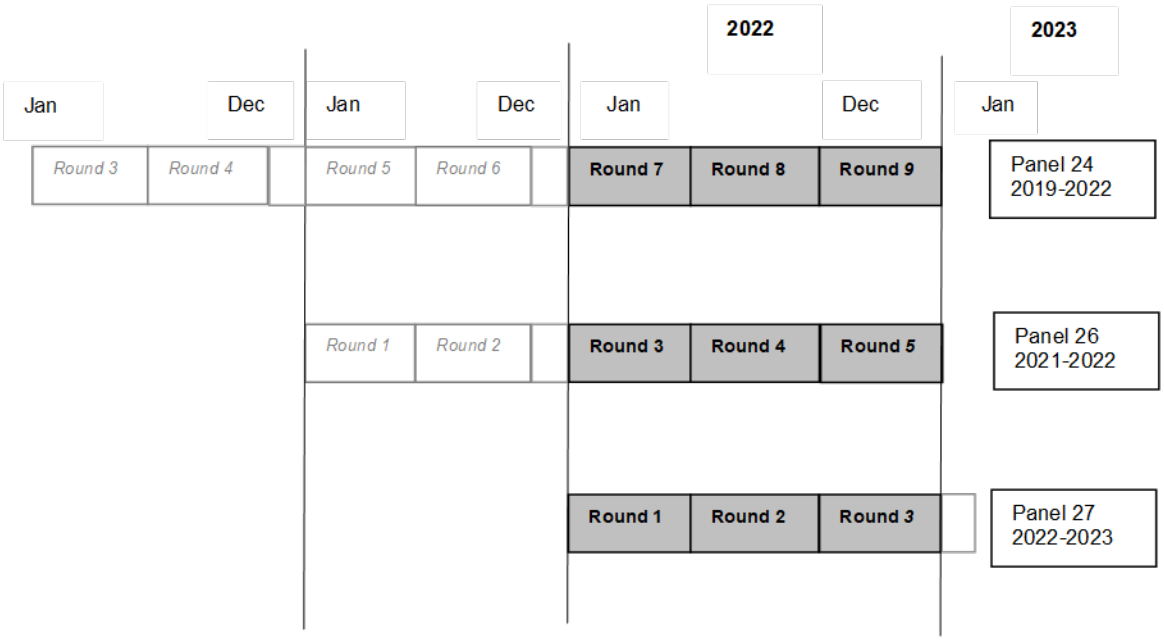
Additional information on MEPS is available from the MEPS project manager or the MEPS public use data manager at the Center for Financing, Access, and Cost Trends, Agency for Healthcare Research and Quality, 5600 Fishers Lane, Rockville, MD 20857 (301-427-1406).

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 data 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 Outpatient Department Visits Public Use File (hereafter referred to as the OP PUF) provides detailed information on outpatient visits for a nationally representative sample of the U.S. civilian noninstitutionalized population and can be used to make estimates of outpatient utilization and expenditures for calendar year 2022. The file contains 57 variables and has a logical record length of 308 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, 2, and the 2022 portion of Round 3 for Panel 27 (i.e., the rounds for the MEPS panels covering calendar year 2022).

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



Each record in this event file represents a unique outpatient event; that is, an outpatient event reported by the household respondent. Outpatient events reported in Panel 24 Round 9, Panel 26 Round 5, and Panel 27 Round 3 and known to have occurred after December 31, 2022 are not included in this file.

Annual counts of outpatient visits 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 PUF can be merged with other MEPS HC data files, for purposes of appending person characteristics such as demographic or health insurance characteristics to each outpatient visit record.

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

This documentation offers a brief overview of the types and levels of data provided, and the content and structure of the file 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)

Any variables not found in this PUF but released on previous years' files may have been excluded because they contained only missing data.

For more information on the MEPS HC sample design, see Chowdhury et al. (2019). For information on the MEPS MPC design, see RTI (2019). Copies of the HC and the MPC survey instruments used to collect the information in the OP PUF are available in the *Survey Questionnaires* section of the MEPS website.

2.0 Data File Information

The 2022 OP PUF consists of one event-level file. The PUF contains characteristics associated with the outpatient (OP) event and imputed expenditure data.

The 2022 OP PUF contains 23,349 outpatient event records; of these records, 22,962 are associated with persons having a positive person-level weight (PERWT22F). This PUF includes outpatient event records for all household members who resided in eligible responding households and for whom at least one outpatient event was reported. Questions inquired whether someone in the family had a visit to an independent lab or testing facility for x-rays or other tests. An affirmative answer to these questions leads to the creation of an office-based provider event record or an outpatient department event record.

Each record represents one household-reported outpatient event that occurred during calendar year 2022. Outpatient visits known to have occurred after December 31, 2022 are not included in this PUF. Some household members may have multiple outpatient events and thus will be represented in multiple records in this PUF. Other household members may have had no outpatient events reported and thus will have no records in this PUF. These data were collected during 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; as well as Rounds 1, 2, and the 2022 portion of Round 3 for Panel 27 of the MEPS HC. The persons represented in this file 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 in the 2022 Consolidated PUF.

Persons with no outpatient visit events for 2022 are not included in this OP PUF but are represented in the person-level 2022 Full Year Population Characteristics PUF (hereafter referred to as the Population Characteristics PUF).

Each outpatient visit record includes the following information: date of the visit; whether the household member saw the doctor; type of care received; type of services (i.e., lab test, sonogram or ultrasound, x-rays, etc.) received; medicines prescribed during the visit; flat fee information; imputed sources of payment; total payment and total charge; a full-year person-level weight; variance strata; and variance PSU.

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 (i.e., Population Characteristics PUF or Consolidated PUF) using the person identifier, DUPERSID. Outpatient visit events in this PUF can also be linked to the MEPS 2022 Medical Conditions Public Use File (hereafter referred to as the Conditions PUF). Please see Section 5.0 for details on how to merge MEPS data files.

2.1 Codebook Structure

For most variables in the OP 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 identifiers

- Unique outpatient visit identifiers
- Outpatient characteristic variables
- Imputed expenditure variables
- Weight and variance estimation variables

Note that the person identifier corresponds to a unique person and the outpatient event identifier corresponds to a unique event.

2.2 Reserved Codes

This OP 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 when there was not enough information from the instrument to calculate the constructed variables. Not having enough information is often the result of skip patterns in the data or of missing information stemming from the responses Refused (-7) or Don't Know (-8). Note that, in addition to Don't Know, reserved code -8 also 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

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

Table 2

Programming Identifiers for Each Variable in the OP PUF

Identifier	Description
Name	Variable name
Description	Variable descriptor
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. All 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 OP PUF were derived from CAPI or the MPC data collection instrument, or were 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";

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:

   - FF - Flat Fee section
   - CP- Charge Payment section
   - OP - Outpatient section
   - TH - Telehealth 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 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

For expenditure variables in the OP PUF, the third character indicates whether the expenditure (or amount paid) 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

MR - Medicare

MD - Medicaid

PV - private insurance

VA - Veterans Administration/CHAMPVA

TR - TRICARE

OF - other federal government

SL - state/local government

WC - Workers’ Compensation

OT - other insurance

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 (22). The eighth character being “X”, indicates whether the variable is edited/imputed.

For example, OPFSF22X is the edited/imputed amount paid by self or family for the facility portion of the expenditure associated with an outpatient visit.

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 ID 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 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 Identifiers (EVNTIDX, FFEEIDX)

EVNTIDX uniquely identifies each outpatient event (i.e., each record in this PUF) and is the variable required to link outpatient events to the 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.

FFEEIDX is a constructed variable that uniquely identifies a flat fee group, that is, all events that were part of a flat fee payment. For example, if a patient receives stitches during an outpatient visit and comes back to have the stitches removed ten days later in a follow-up outpatient visit, both visits are covered under one flat fee dollar amount. These two events (the initial outpatient visit and the subsequent outpatient visit) would have the same value for FFEEIDX. A "mixed" flat fee group could contain both outpatient and office-based visits. Only outpatient and office-based events are allowed in a mixed bundle. Please note that FFEEIDX should be used to link the outpatient and office-based events in order to determine the full set of events that are part of a flat fee group.

Round Indicator (EVENTRN)

EVENTRN indicates the round in which the outpatient event was reported. Please note: Rounds 7 (partial), 8, and 9 are associated with 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 in 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 MPC Data Indicator (MPCDATA)

MPCDATA is a constructed variable that indicates whether MPC data were collected for the outpatient visit. While all outpatient events are sampled into the MPC, not all outpatient event records have MPC data associated with them. This depends upon the cooperation of the household respondent to provide permission forms to contact the outpatient facility as well as the cooperation of the outpatient facility to participate in the survey.

2.5.3 Outpatient Visit Event Variables

This PUF contains variables describing outpatient events reported by respondents in the Outpatient Department section of the MEPS HC questionnaire. The questionnaire contains specific probes for determining details about the outpatient visit. These variables have not been edited.

Visit Details (OPDATEYR-VSTRELCN_M18)

When a person reported a visit to a hospital outpatient department or special clinic, the year and month of the outpatient visit (OPDATEYR and OPDATEMM) and whether the person saw or spoke to a medical doctor (SEEDOC_M18) were ascertained. If the person did not see a specialty doctor (DRSPLTY_M18), or, if the person did not see a physician (i.e., medical doctor), the respondent was asked to identify the type of medical person who was seen (MEDPTYPE_M18). The type of care the person received (VSTCTGRY), and whether the visit was related to a specific condition (VSTRELCN_M18) were also determined. Note that response categories with small frequencies may have been recoded to other categories for confidentiality reasons.

Services, Procedures, and Prescription Medicines (LABTEST_M18-MEDPRESC)

Services received during the visit included whether the person received lab tests (LABTEST_M18), a sonogram or ultrasound (SONOGRAM_M18), x-rays (XRAYS_M18), a mammogram (MAMMOG_M18), an MRI or CAT scan (MRI_M18), an electrocardiogram / an electroencephalogram (EKG_M18), and a vaccination (RCVVAC_M18). Minimal editing was done across treatments, services, and procedures to ensure consistency across -1 (Inapplicable), -7 (Refused), -8 (Don’t Know), and 95 (No Services Received) values. Due to design changes, beginning in 2017, EEG was combined with EKG. Whether a surgical procedure was performed

during the visit was asked (SURGPROC). All the service and procedure variables are set to -1 (Inapplicable) for telehealth events.

Finally, the questionnaire determined if a medicine was prescribed for the person during the visit (MEDPRESC). For a repeat visit event group, if a prescribed medicine is linked to the stem event (MEDPRESC=1), then the value of MEDPRESC is copied to the leaf events without linking the leaf events to the prescribed medicine. MEDPRESC=1 was recoded to -15 (Cannot Be Computed) for all leaf events.

Telehealth (TELEHEALTHFLAG-VISITTYPE)

The Telehealth (TH) module is asked of all events tagged as TH events by the respondent. As part of the TH module, a question asks whether the provider or facility is owned or operated by a hospital. Post-collection, the response to this question is used to reclassify all TH events as either OB or OP. The TH module items were designed to align with the existing OB and OP items to easily allow for reclassifying the event type. All events initially reported as TH also have a categorical variable, VISITTYPE, which indicates whether the visit was over the phone, through real-time video, or some other way.

2.5.4 Clinical Classification Software Refined

Information on household-reported medical conditions (ICD-10-CM condition codes) and aggregated clinically meaningful categories generated using Clinical Classification Software Refined (CCSR) for each outpatient visit are not provided in this file. For information on ICD-10-CM condition codes and associated CCSR codes, see the Conditions PUF.

2.5.5 Flat Fee Variables (FFEEIDX, FFOPTYPE, FFBEF22, FFTOT23)

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: an obstetrician's fee covering a normal delivery, as well as pre- and post-natal care; or a surgeon's fee covering surgical procedure along with 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 in this file include flat fee groups where at least one of the health care events, as reported by the HC respondent, occurred during 2022. By definition a flat fee group can span multiple years. Furthermore, a single person can have multiple flat fee groups.

It is important to note that certain flat fee bundle types reported by household respondents were identified as having a high likelihood of being simple events misidentified as bundle events. To address this, starting in 2022, HC-reported flat fee bundles were considered as flat fees if the bundle consisted only of dental events, or the bundle started in the previous year and also had events in 2022.

Other HC-reported bundles were not allowed as flat fee bundles, and events in these bundles were treated as simple events. HC-reported bundles that included a mix of emergency room and hospitalization events were treated as linked events. All emergency room expenditures were combined with hospital inpatient expenditures. However, provider-reported flat fees were processed in a similar way to prior years.

Flat Fee Variable Descriptions

Flat Fee ID (FFEEIDX)

As noted in "Record Identifiers," the variable FFEEIDX uniquely identifies all events that are part of the same flat fee group for a person. On any 2022 MEPS event file, every event that was a part of a specific flat fee group has the same value for FFEEIDX. Note that prescribed medicine and home health events are never included in a flat fee group and none of the flat fee variables are not a variable in those event PUFs.

Flat Fee Type (FFOPTYPE)

FFOPTYPE indicates whether the 2022 outpatient visit is the "stem" or "leaf" of a flat fee group. A stem (records with FFOPTYPE = 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 FFOPTYPE = 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 outpatient visits that are not part of a flat fee payment, the FFOPTYPE is set to Inapplicable (-1).

Counts of Flat Fee Events that Cross Years (FFBEF22, FFTOT23)

As described in "Definition of Flat Fee Payments", a flat fee payment covers multiple events and the multiple events could span multiple years. For situations where the outpatient visit occurred in 2022 as a part of a group of events, and some of the events occurred before or after 2022, counts of the known events are provided on the outpatient visit record. Variables that indicate events that that occurred before or after 2022 are as follows:

> FFBEF22 - total number of pre-2022 events in the same flat fee group as the 2022 outpatient visit. This count would not include the 2022 outpatient visit(s).
>
> FFTOT23 - the number of 2023 outpatient visits expected to be in the same flat fee group as the outpatient visit record that occurred in 2022.

If there are no 2021 events in this PUF, FFBEF22 will be omitted. Likewise, if there are no 2023 events in this PUF, FFTOT23 will be omitted. If there are no flat fee data related to the records in this PUF, FFEEIDX and FFOPTYPE will be omitted as well. Please note that the Variable-Source Crosswalk in this document lists all possible flat fee variables.

Caveats of Flat Fee Groups

There are 204 outpatient visits that are identified as 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 2022 but the remaining visits that were part of this flat fee group occurred in 2023. In this case, the 2022 flat fee group represented in this file would consist of one event (the stem). The 2023 leaf 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 2021 but subsequent visits occurred during 2022. In this case, the initial visit would not be represented on the file. This 2022 flat fee group would then only consist of one or more leaf records and no stem. Another reason for which a flat fee group would not have a stem and at least one leaf record is that the stem or leaves could have been reported as different event types. Outpatient and office-based medical provider visits are the only two event types allowed in a single flat fee group. The stem may have been reported as an outpatient department visit and the leaves may have been reported as office-based medical provider visits. Note that the Variable-Source Crosswalk in this document lists all possible flat fee variables.

2.5.6 Expenditure Data

Definition of Expenditures

Expenditures in this PUF refer to payments for outpatient services. More specifically, expenditures in MEPS are defined as the sum of payments for care received for each outpatient visit, 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, the estimates do not incorporate any payment not directly tied to specific medical care visits, such as bonuses or retrospective payment adjustments paid by third party payers. Currently, 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. 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.

Expenditure data related to outpatient visits are broken out by facility and separately billing doctor expenditures. When a facility bills directly for the services provided by physicians and other providers, in MEPS, the facility charge and payments include the physician and other providers' charge and payments. This file contains six categories of expenditure variables per visit: basic hospital outpatient facility expenses; expenses for doctors who billed separately from the outpatient facility for any services provided during the outpatient visit; 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, please refer to Section 3.5 for more information.

Data Editing and Imputation Methodologies of Expenditure Variables

The expenditure data included in this PUF were derived from both the MEPS HC and the 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 outpatient visits, MPC data were used if available; otherwise, HC data were used. Missing data for outpatient visits where HC data were not complete and MPC data were not collected, or MPC data were not complete, were derived through the imputation process.

General Data Editing Methodology

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

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 missing payment was used to impute the missing payment value. The weighted sequential hot-deck procedure was used to impute the missing total charges. This procedure uses survey data from donors to replace missing data while taking into account the donors' weighted distribution in the imputation process to ensure that the weighted distribution of recipients' expenditures reflects the weighted distribution of the donors' expenditures. The imputations for the flat fee events were carried out separately from the simple events.

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.

Outpatient Visit Data Editing and Imputation

Facility expenditures for outpatient services 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 in the reporting of expenditures. After the data from each source were edited, a decision was made whether to use household- or MPC-reported information in the final editing and predictive mean matching imputations for missing expenditures. The general rule was that MPC data would be used where a household-reported event corresponded to an

MPC-reported event (i.e., a matched event), since providers usually have more complete and accurate data on sources and amounts of payment than households.

One of the more important edits separated flat fee events from simple events. This edit was necessary because groups of events covered by a flat fee (i.e., a flat fee bundle) were edited and imputed separately from individual events covered by a single charge (i.e., simple events). (See Section 2.5.5 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 predictive mean matching 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 separate recipient categories.

The logical edits produced eight recipient categories in which all events had a common extent of missing data. However, for predictive mean matching imputations, the recipients were grouped into four categories based on the known status of total charge and the sources of payment: (1) known charge but unknown payment status of at least one potential paying source; (2) unknown charge and unknown payment status of at least one potential paying source; (3) known charge and known status of all payment sources; and (4) unknown charge and known status of all payment sources. Separate predictive mean matching imputations were performed on events in each recipient group. For outpatient events, the donor pool was restricted to events with complete expenditures from the MPC. To improve the reliability of imputation, current year donors and inflation-adjusted prior year donors are used for the predictive mean matching imputations.

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).

For office-based and outpatient events, the donor pool also included events originally reported by providers as paid on a capitated basis. To obtain the fee-for-service (FFS) equivalent payments for these capitated events, a “capitation imputation” was implemented (see below). Once imputed with the FFS equivalent payments, these events became donors for all other incomplete events, particularly for events reported by the household as services covered under managed care plans.

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.

Capitation Imputation

The weighted sequential hot-deck procedure was used to estimate expenditures at the event-level for events that were paid on a per-month per-person (capitated) basis. The capitation imputation procedure was designed as a reasonable approach to complete event-level expenditures for persons in non-fee for service managed care plans. HMO events reported in the MPC as covered by capitation arrangements were imputed using similar HMO events paid on a fee-for-service, with total charge as a key variable. Then this fully completed set of MPC events was used in the donor pool for the main imputation process for cases in HMOs. By using this strategy, capitated HMO events were imputed as if the provider were reimbursed from the HMO on a discounted fee-for-service basis.

Imputation Flag (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 OB 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

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 still be present. Thus, if the first visit in the flat fee group occurred prior to 2022, all of the events that occurred in 2022 will have zero payments. Conversely, if the first event in the flat fee group occurred at the end of 2022, the total expenditure for the entire flat fee group will be on that event, regardless of the number of events it covered after 2022. See Section 2.5.5 for details on the flat fee variables.

Zero Expenditures

There are some medical events reported by respondents where the payments were zero. Zero payment events can occur in MEPS for the following reasons: (1) the visit was covered under a flat fee arrangement (flat fee payments are included only on the first event covered by the

arrangement), (2) there was no charge for a follow-up visit, (3) the provider was never paid directly for services provided by an individual, insurance plan, or other source, (4) the charges were included in another bill, or (5) the event was paid through government or privately funded research or clinical trials.

Discount Adjustment Factor

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

Sources of Payment

In addition to total expenditures, variables are provided which itemize expenditures according to major 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.

Prior to 2019, for cases where reported insurance coverage and sources of payment are inconsistent, the positive amount from a source inconsistent with reported insurance coverage was moved to one or both of the source categories Other Private and Other Public. Beginning in 2019, this step is removed and the inconsistency between the payment sources and insurance coverage is allowed to remain - the amounts are not moved to Other Private and Other Public categories any more. The two source of payment categories, Other Private and Other Public, are no longer available.

Imputed Outpatient Expenditure Variables

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

Outpatient Facility Expenditure Variables (OPFSF22X-OPFOT22X, OPFXP22X, OPFTC22X)

Outpatient visit expenses include all expenses for treatment, services, tests, diagnostic and laboratory work, x-rays, and similar charges, as well as any physician services included in the hospital outpatient visit charge.

OPFSF22X - OPFOT22X are the 10 sources of payment. The 10 sources of payment are: self/family (OPFSF22X), Medicare (OPFMR22X), Medicaid (OPFMD22X), private insurance (OPFPV22X), Veterans Administration/CHAMPVA (OPFVA22X), TRICARE (OPFTR22X), other federal sources (OPFOF22X), state and local (non-federal) government sources (OPFSL22X), Workers' Compensation (OPFWC22X), and other insurance (OPFOT22X). OPFXP22X is the sum of the 10 sources of payment for the outpatient facility expenditures, and OPFTC22X is the total charge. Please note that where an outpatient visit record is linked to a hospital inpatient stay record, all facility sources of payment variables, as well as OPFTC22X have been zeroed out.

Outpatient Physician Expenditures (OPDSF22X - OPDOT22X, OPDXP22X, OPDTC22X)

Charges for services provided in a hospital setting by physicians and other providers are sometimes billed directly by the hospital. In such cases, these charges are included in the hospital-facility charge and payments. When the charges are not billed directly by the hospital, physicians and other providers bill charges for the provided services directly to the third-party and the patient. In such cases, these providers are called separately billing doctors (SBD). 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 the outpatient facility bill.

For physicians who bill separately (i.e., outside the outpatient facility bill), a separate data collection effort within the MPC was performed to obtain the 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, an outpatient visit could have a radiologist and a pathologist associated with it. If their services are not included in the outpatient visit bill then this is one medical event with 2 separately billing doctors. The imputed expenditure information associated with the separately billing doctors was summed to the event-level and is provided on the file. OPDSF22X - OPDOT22X are the 10 sources of payment,

OPDXP22X is the sum of the 10 sources of payments, and OPDTC22X is the physician(s) total charge.

Analysts should consider 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).

Total Expenditures and Charges for Outpatient Visits (OPXP22X, OPTC22X)

Analysts interested in total expenditures should use the variable OPXP22X, which includes both facility and physician amounts. Those interested in total charges should use the variable OPTC22X, which includes both facility and physician charges (see "Definition of Expenditures" for an explanation of the "charge" concept).

2.5.7 Rounding

Expenditure variables in the OP PUF have been rounded to the nearest penny. Person-level expenditure information to be released on the Consolidated PUF were rounded to the nearest dollar. It should be noted that using the MEPS 2022 event files to create person-level totals will yield slightly different totals than those found 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 Consolidated 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 on 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 file can be used to develop national 2022 event-level estimates for the U.S. civilian noninstitutionalized population on outpatient visits as well as expenditures, and sources of payment for these visits. Estimates of total visits are the sum of the weight variable (PERWT22F) across relevant event records while estimates of other variables must be weighted by PERWT22F to be nationally representative. The tables below contain event-level estimates for selected variables.

Table 3

Selected Event-Level Estimates - Outpatient Visits

Estimate of Interest	Variable Name	Estimate (SE)	Estimate Excluding Zero Payment Events (SE)*
Total number of outpatient visits (in millions)	PERWT22F	294.9 (17.56)	289.6 (17.23)
Total number of in-person visits to doctor (SEEDOC_M18=1, in millions)	PERWT22F	140.6 (9.02)	138.2 (8.92)
Proportion of outpatient visits with expenditures > 0*	OPXP22X	0.982 (0.0025)	----------

Table 4

Selected Event-Level Estimates - Outpatient Expenditures

Estimate of Interest	Variable Name	Estimate (SE)	Estimate Excluding Zero Payment Events (SE)*
Mean total payments per visit (all sources)	OPXP22X	$904 ($39.4)	$920 ($40.1)
Mean out-of-pocket payment per visit	OPDSF22X +OPFSF22X	$76 ($6.0)	$78 ($6.1)
Mean proportion of total expenditures paid by private insurance per visit	(OPDPV22X +OPFPV22X) /OPXP22X	----------	0.361 (0.0153)

Table 5

Selected Event-Level Estimates - Expenditures: Physician Visits

Estimate of Interest	Variable Name	Estimate (SE)	Estimate Excluding Zero Payment Events (SE)*
Mean total payments per visit where person saw medical doctor	OPXP22X	$1,321 ($73.7)	$1,344 ($75.4)

Estimate of Interest	Variable Name	Estimate (SE)	Estimate Excluding Zero Payment Events (SE)*
Mean out-of-pocket payment per visit where person saw medical doctor	OPDSF22X +OPFSF22X	$113 ($10.6)	$115 ($10.9)
Mean proportion of total expenditures per visit paid by private insurance where person saw medical doctor	(OPDPV22X +OPFPV22X) /OPXP22X	----------	0.377 (0.0171)

* Zero payment events can occur in MEPS for the following reasons: (1) the visit was covered under a flat fee arrangement (flat fee payments are included only on the first event covered by the arrangement), (2) there was no charge for a follow-up visit, (3) the provider was never paid directly for services provided by an individual, insurance plan, or other source, (4) the charges were included in another bill, or (5) the event was paid through government or privately funded research or clinical trials.

4.2 Person-Based Estimates for Outpatient Visits

To enhance analyses of hospital outpatient visits, analysts may link information about outpatient visits by sample persons in this file to the annual 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 (see Section 5 below for more details). Both this PUF and the Consolidated PUF may be used to derive estimates for persons with outpatient care and annual estimates of total expenditures. However, for estimates that pertain to those who did not have hospital outpatient care as well as those who did (for example, the percentage of adults with at least one outpatient event during the past year or the mean number of outpatient events in the past year among those 65 or older), this file cannot be used. Only those persons with at least one outpatient event are represented in this PUF. The Consolidated PUF must be used for person-level analyses that include both persons with and without hospital outpatient 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, flat fee, and zero expenditures) are described in “Data Editing and Imputation Methodologies of Expenditure Variables.”

4.4 Variance Estimation (VARSTR, VARPSU)

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 in this file as well as all other MEPS PUFs. Software packages that permit the use of the Taylor-series linearization method include SUDAAN, R, Stata, 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 in this OP 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 consisted 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 are based on the same NHIS design, there are only 165 variance strata. These 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 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 2001-2117, 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 PUFs, 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 in this MEPS PUF for the purposes of variance estimation. However, a file containing a BRR replication structure is made available so that 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 nonlinear 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.*

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. & J.W. Cohen. (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

MEPS HC 239F: 2022 OUTPATIENT DEPARTMENT 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
FFEEIDX	Flat Fee ID	CAPI derived
MPCDATA	MPC data flag	Constructed

Outpatient Department Visit Variables

Variable	Description	Source
OPDATEYR	Event date - year	CAPI derived
OPDATEMM	Event date - month	CAPI derived
SEEDOC_M18	Did person talk to MD this visit	OP10
DRSPLTY_M18	OPAT doctor's specialty	OP20
MEDPTYPE_M18	Type of medical P talked to on visit date	OP30
VSTCTGRY	Best category for care P received on visit date	OP40
VSTRELCN_M18	This visit/phone call related to spec condition	OP50
LABTEST_M18	This visit did P have lab tests	OP80
SONOGRAM_M18	This visit did P have sonogram or ultrasound	OP80
XRAYS_M18	This visit did P have x-rays	OP80
MAMMOG_M18	This visit did P have a mammogram	OP80
MRI_M18	This visit did P have an MRI/Catscan	OP80
EKG_M18	This visit did P have an EKG, EEG or ECG	OP80
RCVVAC_M18	This visit did P receive a vaccination	OP80
SURGPROC	Was surgical procedure performed on P this visit	OP70
MEDPRESC	Any medicine prescribed for P during visit	OP90

Variable	Description	Source
TELEHEALTHFLAG	Is this a telehealth event	Constructed
VISITTYPE	Type of telehealth visit	TH10

Flat Fee Variables

Variable	Description	Source
FFOPTYPE	Flat fee bundle	Constructed
FFBEF22	Total # of visits in FF before 2022	FF50
FFTOT23	Total # of visits in FF after 2022	FF60

Imputed Total Expenditure Variables

Variable	Description	Source
OPXP22X	Total expenditure for event (OPFXP22X+OPDXP22X)	Constructed
OPTC22X	Total charge for event (OPFTC22X+OPDTC22X)	Constructed

Imputed Facility Expenditure Variables

Variable	Description	Source
OPFSF22X	Facility amount paid, self/family (Imputed)	CP Section (Edited)
OPFMR22X	Facility amount paid, Medicare (Imputed)	CP Section (Edited)
OPFMD22X	Facility amount paid, Medicaid (Imputed)	CP Section (Edited)
OPFPV22X	Facility amount paid, private insurance (Imputed)	CP Section (Edited)
OPFVA22X	Facility amount paid, Veterans/CHAMPVA (Imputed)	CP Section (Edited)
OPFTR22X	Facility amount paid, TRICARE (Imputed)	CP Section (Edited)
OPFOF22X	Facility amount paid, other federal (Imputed)	CP Section (Edited)
OPFSL22X	Facility amount paid, state & local government (Imputed)	CP Section (Edited)

Variable	Description	Source
OPFWC22X	Facility amount paid, workers' compensation (Imputed)	CP Section (Edited)
OPFOT22X	Facility amount paid, other insurance (Imputed)	CP Section (Edited)
OPFXP22X	Facility sum payments OPFSF22X - OPFOT22X	Constructed
OPFTC22X	Total facility charge (Imputed)	CP Section (Edited)

Imputed Physician Expenditure Variables

Variable	Description	Source
OPDSF22X	Doctor amount paid, self/family (Imputed)	Constructed
OPDMR22X	Doctor amount paid, Medicare (Imputed)	Constructed
OPDMD22X	Doctor amount paid, Medicaid (Imputed)	Constructed
OPDPV22X	Doctor amount paid, private insurance (Imputed)	Constructed
OPDVA22X	Doctor amount paid, Veterans/CHAMPVA (Imputed)	Constructed
OPDTR22X	Doctor amount paid, TRICARE (Imputed)	Constructed
OPDOF22X	Doctor amount paid, other federal (Imputed)	Constructed
OPDSL22X	Doctor amount paid, state & local government (Imputed)	Constructed
OPDWC22X	Doctor amount paid, workers' compensation (Imputed)	Constructed
OPDOT22X	Doctor amount paid, other insurance (Imputed)	Constructed
OPDXP22X	Doctor sum payments OPDSF22X - OPDOT22X	Constructed
OPDTC22X	Total doctor charge (Imputed)	Constructed
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