

# **MEPS HC-206B: 2018 Dental Visits**

## **May 2020**

---

The MEPS instrument design changed beginning in Spring of 2018, affecting Panel 23 Round 1, Panel 22 Round 3, and Panel 21 Round 5. For the Full-Year 2017 PUFs, the Panel 22 Round 3 and Panel 21 Round 5 data were transformed to the degree possible to conform to the previous design.

The Full-Year 2018 PUFs are the first year all rounds of data were collected with the re-designed instrument, and no data were transformed to conform to the previous design. In addition, the value -9 NOT ASCERTAINED was removed as an allowable value in the Full-Year 2018 PUFs. **Data users should be aware of possible impacts on the data and especially trend analysis for these data years due to the design transition.**

---

**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.....		A-1
B. BACKGROUND .....		B-1
1.0 Household Component .....		B-1
2.0 Medical Provider Component.....		B-1
3.0 Survey Management and Data Collection .....		B-2
C. TECHNICAL AND PROGRAMMING INFORMATION .....		C-1
1.0 General Information .....		C-1
2.0 Data File Information .....		C-2
2.1 Codebook Structure.....		C-3
2.2 Reserved Codes.....		C-3
2.3 Codebook Format.....		C-3
2.4 Variable Source and Naming Conventions.....		C-4
2.4.1 Variable - Source Crosswalk.....		C-4
2.4.2 Expenditure and Source of Payment Variables.....		C-4
2.5 File Contents .....		C-5
2.5.1 Survey Administration Variables .....		C-5
2.5.2 Dental Event Variables .....		C-6
2.5.3 Flat Fee Variables (FFEEIDX, FFDVTYPE, FFBEFC18, FFTOT19).....		C-7
2.5.4 Condition Codes.....		C-8
2.5.5 Expenditure Data.....		C-8
3.0 Sample Weight (PERWT18F) .....		C-13
3.1 Overview.....		C-13
3.2 Details on Person Weight Construction.....		C-13
3.2.2 MEPS Panel 23 Weight Development Process...		C-14
3.2.3 The Final Weight for 2018.....		C-15
3.2.4 Coverage .....		C-15
3.3 Using MEPS Data for Trend Analysis .....		C-15

## Table of Contents (continued)

4.0	Strategies for Estimation.....	C-16
4.1	Developing Event-Level Estimates .....	C-16
4.2	Person-Based Estimates for Dental Care.....	C-18
4.3	Variables with Missing Values .....	C-18
4.4	Variance Estimation (VARPSU, VARSTR).....	C-18
4.4.1	Taylor-series Linearization Method .....	C-19
4.4.2	Balanced Repeated Replication (BRR) Method..	C-20
5.0	Merging/Linking MEPS Data Files.....	C-21
5.1	Linking to the Person-Level File.....	C-21
5.2	Linking to the Prescribed Medicines File.....	C-22
5.3	Linking to the Medical Conditions File.....	C-22
D.	VARIABLE-SOURCE CROSSWALK.....	D-1
	VARIABLE-SOURCE CROSSWALK.....	D-1

## **A. Data Use Agreement**

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

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

1. No one is to use the data in this data set in any way except for statistical reporting and analysis; and
2. If the identity of any person or establishment should be discovered inadvertently, then (a) no use will be made of this knowledge, (b) the Director Office of Management AHRQ will be advised of this incident, (c) the information that would identify any individual or establishment will be safeguarded or destroyed, as requested by AHRQ, and (d) no one else will be informed of the discovered identity; and
3. No one will attempt to link this data set with individually identifiable records from any data sets other than the Medical Expenditure Panel Survey or the National Health Interview Survey. 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 health care. Estimates can be produced for individuals, families, and selected population subgroups. The panel design of the survey, which includes 5 Rounds of interviews covering 2 full calendar years, provides data for examining person level changes in selected variables such as expenditures, health insurance coverage, and health status. Using computer assisted personal interviewing (CAPI) technology, information about each household member is collected, 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. Each annual MEPS-HC sample size is about 15,000 households. Data can be analyzed at either the person or 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 NHIS implemented a new sample design, which included Asian persons in addition to households with Black and Hispanic persons in the oversampling of minority populations. NHIS introduced a new sample design in 2016 that discontinued oversampling of these minority groups. The linkage of the MEPS to the previous year's NHIS provides additional data for longitudinal analytic purposes.

### **2.0 Medical Provider Component**

Upon completion of the household CAPI interview and obtaining permission from the household survey respondents, a sample of medical providers are 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 information is collected on dates of visits, diagnosis and procedure codes, charges and payments. The Pharmacy Component (PC), a subcomponent of the MPC, does not collect charges or diagnosis and procedure codes but does collect drug detail information, including National Drug Code (NDC) and medicine name, as well as date filled and sources and 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. Data are collected under contract with Westat, Inc. (MEPS HC) and Research Triangle Institute (MEPS MPC). Data sets and summary statistics are edited and published in accordance with the confidentiality provisions of the Public Health Service Act and the Privacy Act. The National Center for Health Statistics (NCHS) provides consultation and technical assistance.

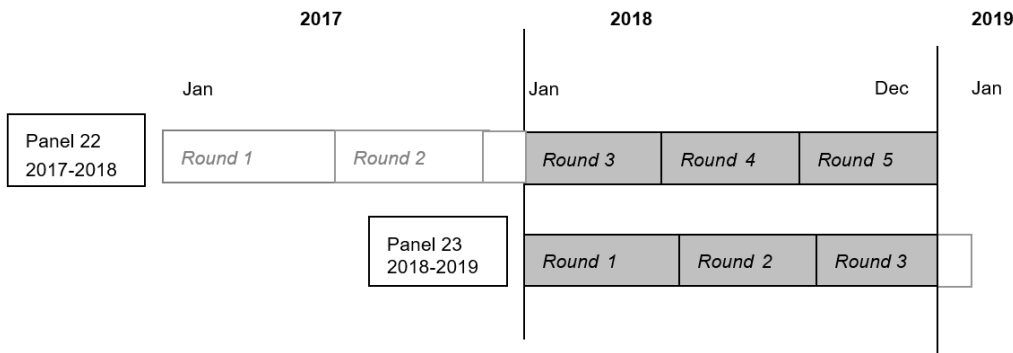
As soon as data collection and editing are completed, the MEPS survey data are released to the public in staged releases of summary reports, micro data files, and tables via the [MEPS website](#). Selected data can be analyzed through MEPSnet, an on-line interactive tool designed to give data users the capability to statistically analyze MEPS data in a menu-driven environment.

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 2018 Medical Expenditure Panel Survey (MEPS) Household Component (HC). Released as an ASCII data file (with related SAS, SPSS, and Stata programming statements) and a SAS transport file, the 2018 Dental public use file provides detailed information on dental events for a nationally representative sample of the civilian noninstitutionalized population of the United States. Data from the Dental file can be used to make estimates of dental event utilization and expenditures for calendar year 2018. The file contains 60 variables and has a logical record length of 270 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 2018 portion of Round 3 and Rounds 4 and 5 for Panel 22, as well as Rounds 1, 2 and the 2018 portion of Round 3 for Panel 23 (i.e., the rounds for the MEPS panels covering calendar year 2018).



Each record on this event file represents a unique dental event; that is, a dental event reported by the household respondent. Counts of dental event utilization are based entirely on household reports. Dental events were not included in the Medical Provider Component (MPC); therefore, all expenditure and payment data on the Dental event file are reported by the household.

Data from this event file can be merged with other 2018 MEPS HC data files for the purposes of appending person-level data such as demographic characteristics or health insurance coverage to each dental record.

This file can also be used to construct summary variables of expenditures, sources of payment, and related aspects of the dental event. Aggregate annual person-level information on the use of dental events and other health services is provided on the MEPS 2018 Full Year Consolidated Data File 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 file and the codebook. It contains the following sections:

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

For more information on the MEPS HC sample design, see Chowdhury et al (2019). A copy of the MEPS HC survey instrument used to collect the information on the Dental file is available on the [MEPS website](#).

## **2.0 Data File Information**

The 2018 Dental public use data set consists of one event-level data file. The file contains characteristics associated with the dental event and imputed expenditure data.

The 2018 Dental public use data set contains 29,583 dental event records; of these records, 29,184 are associated with persons having a positive person-level weight (PERWT18F). This file includes dental event (DV) records for all household members who resided in eligible responding households and reported at least one dental event. Each record represents one household-reported dental event that occurred during calendar year 2018. Dental visits known to have occurred before January 1, 2018 and after December 31, 2018 are not included on this file. Some household members may have multiple dental events and thus will be represented in multiple records on this file. Other household members may have had reported no dental events and thus will have no records on this file. These data were collected during the 2018 portion of Round 3, and Rounds 4 and 5 for Panel 22, as well as Rounds 1, 2, and the 2018 portion of Round 3 for Panel 23 of the MEPS HC. The persons represented on this file had to meet either (a) or (b) below:

- a) Be classified as a key in-scope person who responded for his or her entire period of 2018 eligibility (i.e., persons with a positive 2018 full-year person-level weight (PERWT18F > 0)), or
- b) Be an eligible member of a family all of whose key in-scope members have a positive person-level weight (PERWT18F > 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 (FAMWT18F > 0). Note that FAMIDYR and FAMWT18F are variables on the 2018 Full Year Consolidated Data File.

Persons with no dental events for 2018 are not included on this event-level DV file but are represented on the person-level 2018 Full Year Population Characteristics file.

Each dental event record includes the following: date of the dental event; type of provider seen; procedure(s) associated with the dental event; flat fee information; imputed sources of payment; total payment and total charge of the dental 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 2018 MEPS HC person-level data (e.g. Full Year Consolidated or Full Year Population Characteristics files) using the person identifier, DUPERSID. Dental events can also be linked to the MEPS 2018 Prescribed Medicines File. Please see Section 5.0 or the 2018 Appendix for details on how to merge MEPS data files.

## 2.1 Codebook Structure

For most variables on the Dental event file, 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 sequence list variables in the following order:

- Unique person identifier
- Unique dental event identifier
- Dental characteristic variables
- Imputed expenditure variables
- Weight and variance estimation variables

Note that the person identifier is unique within this data year.

## 2.2 Reserved Codes

The following reserved code values are used:

Value	Definition
-1 INAPPLICABLE	Question was not asked due to skip pattern
-7 REFUSED	Question was asked and respondent refused to answer question
-8 DK	Question was asked and respondent did not know answer
-15 CANNOT BE COMPUTED	Value cannot be derived from data

Generally, values of -1, -7, -8, and -15 for non-expenditure variables have not been edited on this file. The values of -1 and -15 can be edited by the data users/analysts by following the skip patterns in the [HC survey questionnaire](#) located on the MEPS website.

## 2.3 Codebook Format

The codebook describes an ASCII data set (although the data are also being provided in a SAS transport file).

The following codebook items are provided for each variable:

<b>Identifier</b>	<b>Description</b>
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, variable names reflect the content of the variable. All imputed/edited variables end with an “X”.

Beginning in 2018, as variable collection, universe, or categories are altered, the variable name will be appended with “\_Myy” to indicate in which year the alterations took place. Details about these alterations can be found throughout this document.

### **2.4.1 Variable - Source Crosswalk**

Variables were derived from the HC survey questionnaire or from the CAPI. 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 so indicated as “CAPI derived” or “Assigned in sampling,” respectively;
2. Variables which come from one or more specific questions have those questionnaire sections and question numbers indicated in the “Source” column; questionnaire sections are identified as:

FF - Flat Fee section  
 DN - Dental Event section  
 CP - Charge Payment section

3. Variables constructed from multiple questions using complex algorithms are labeled “Constructed” in the “Source” column; and
4. Variables that have been edited or imputed are so indicated.

### **2.4.2 Expenditure and Source of Payment Variables**

The names of the expenditure and source of payment variables follow a standard convention, are seven characters in length, and end in an “X” indicating edited/imputed. Please note that imputed

means that a series of logical edits, as well as an imputation process to account for missing data, have been performed on the variable.

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

The first two characters indicate the type of event:

IP - inpatient stay	OB - office-based visit
ER - emergency room visit	OP - outpatient visit
HH - home health visit	DV - dental visit
OM - other medical equipment	RX - prescribed medicine

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

SF - self or family	OF - other federal government
MR - Medicare	SL - state/local government
MD - Medicaid	WC - Workers' Compensation
PV - private insurance	OT - other insurance
VA - Veterans Administration/CHAMPVA	OR - other private
TR - TRICARE	OU - other public
	XP - sum of payments

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

The fifth and sixth characters indicate the year (18). The seventh character, "X", indicates the variable is edited/imputed.

For example, DVSF18X is the edited/imputed amount paid by self or family for 2018 dental expenditures. Out-of-pocket by self or family (SF) 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.

## **2.5 File Contents**

### **2.5.1 Survey Administration Variables**

#### **2.5.1.1 Person Identifiers (DUID, PID, DUPERSID)**

The dwelling unit ID (DUID) is a seven-digit random number assigned after the case was sampled for MEPS. The three-digit person number (PID) uniquely identifies each person within the dwelling unit. The ten-character variable DUPERSID uniquely identifies each person represented on the file and is the combination of the variables DUID and PID. As part of the new CAPI design, the lengths of the ID variables have changed in the file. The additional 2 bytes in the IDs resulted from adding a 2-digit panel number to the beginning of all the IDs. Analysts wishing to pool data years 2017 and 2018 should add panel numbers to the beginning of Panel 22 Year 2017 ID variables or remove the 2-digit panel number at the beginning of Panel 22 Year 2018 ID variables to ensure they identify the same person.

For detailed information on dwelling units and families, please refer to the documentation for the 2018 Full Year Population Characteristics File.

### **2.5.1.2 Record Identifiers (EVNTIDX, FFEEIDX)**

EVNTIDX uniquely identifies each dental event (i.e., each record on the Dental file) and is the variable required to link dental events to data files containing details on prescribed medicines (MEPS 2018 Prescribed Medicines File). As part of the new CAPI design, the length of the EVNTIDX has changed to 16 in the file. In addition to the 2-digit panel number added in the beginning, a 2-digit event type number is added to the end. For details on linking see Section 5.0 or the MEPS 2018 Appendix File, HC-206I.

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, a charge for orthodontia is typically covered in a flat fee arrangement where all visits are covered under one flat fee dollar amount. These events would have the same value for FFEEIDX. FFEEIDX identifies a flat fee payment that was identified using information from the Household Component.

### **2.5.1.3 Round Indicator (EVENTRN)**

EVENTRN indicates the round in which the dental event was reported. Please note: Rounds 3 (partial), 4, and 5 are associated with MEPS survey data collected from Panel 22. Likewise, Rounds 1, 2, and 3 (partial) are associated with data collected from Panel 23.

### **2.5.1.4 Panel Indicator (PANEL)**

PANEL is a constructed variable used to specify the panel number for the person. PANEL will indicate either Panel 22 or Panel 23 for each person on the file. Panel 22 is the panel that started in 2017, and Panel 23 is the panel that started in 2018.

## **2.5.2 Dental Event Variables**

This file contains variables describing dental events reported by household respondents in the Dental Section of the MEPS HC survey questionnaire.

### **2.5.2.1 Date of Visit (DVDATEYR – DVDATEMM)**

There are variables that indicate the month and year a dental event occurred (DVDATEMM and DVDATEYR, respectively). These variables have not been edited or imputed.

### **2.5.2.2 Type of Provider Seen (GENDENT\_M18 – DENTYPE\_M18)**

Respondents were asked about the type of provider seen during the dental visit (e.g., general dentist, pediatric dentist, dental hygienist, or specialist). More than one type of provider may have been identified on an event record.

### **2.5.2.3 Treatment, Procedures, and Services (EXAMINE\_M18 – ORTHDONT\_M18)**

Respondents were asked about the types of services or treatments received during the visit (EXAMINE\_M18 – ORTHDONT\_M18), such as root canal or x-rays. More than one type of service or treatment may have been identified on an event record. Some procedures or services identified in DENTOTHX as “Other specify dental procedures edited” have been edited to appropriate procedure and service categories. While the unedited versions of these variables are included in the dental event file every year, an edited version of a particular variable is included only if editing was done for that category. Starting in FY 2015, DENTOTHR was dropped from the file, but the edited version of this variable (DENTOTHX, “Other specify dental procedures edited”) is still included. Please note that the crosswalk in this document lists all possible edited procedure and service category variables; the edited variables in the data file will differ by year.

### **2.5.3 Flat Fee Variables (FFEEIDX, FFDVTYPE, FFBEF18, FFTOT19)**

#### **2.5.3.1 Definition of Flat Fee Payments**

A flat fee is the fixed dollar amount a person is charged for a package of services provided during a defined period of time. Examples would be an orthodontist’s fee, which covers multiple visits; or a dental surgeon’s fee, which covers surgical procedure and post-surgical care. A flat fee group is the set of medical services that are covered under the same flat fee payment. The flat fee groups represented on the dental file include flat fee groups where at least one of the health care events, as reported by the HC respondent, occurred during 2018. By definition, a flat fee group can span multiple years. Furthermore, a single person can have multiple flat fee groups.

#### **2.5.3.2 Flat Fee Variable Descriptions**

##### **2.5.3.2.1 Flat Fee ID (FFEEIDX)**

As noted earlier in Section 2.5.1.2 “Record Identifiers,” the variable FFEEIDX uniquely identifies all events that are part of the same flat fee group for a person. On any 2018 MEPS event file, every event that is part of a specific flat fee group will have the same value for FFEEIDX. Note that prescribed medicine and home health events are never included in a flat fee group and none of the flat fee variables is on those event files.

##### **2.5.3.2.2 Flat Fee Type (FFDVTYPE)**

FFDVTYPE indicates whether the 2018 dental event is the “stem” or “leaf” of a flat fee group. A stem (records with FFDVTYPE = 1) is the initial dental service (event) which is followed by other dental events that are covered under the same flat fee payment. The leaves of the flat fee group (records with FFDVTYPE = 2) are those dental 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 dental visits that are not part of a flat fee payment, the FFDVTYPE is set to -1, “INAPPLICABLE”.

### **2.5.3.2.3 Counts of Flat Fee Events that Cross Years (FFBEF18, FFTOT19)**

As described in Section 2.5.3.1, a flat fee payment covers multiple events and the multiple events could span multiple years. For situations where a 2018 dental visit is part of a group of events, and some of the events occurred before or after 2018, counts of the known events are provided on the dental record. Variables that indicate events occurring before or after 2018 are the following:

FFBEF18 – indicates total number of pre-2018 events in the same flat fee group as the 2018 dental event. This count would not include 2018 dental events.

FFTOT19 – indicates the number of 2019 medical events expected to be in the same flat fee group as the 2018 dental event record.

### **2.5.3.3 Caveats of Flat Fee Groups**

Data users/analysts should note that flat fee payments are common on the dental file. There are 5,371 dental events that are identified as being part of a flat fee payment group. In general, every flat fee group should have an initial visit (stem) and at least one subsequent visit (leaf). There are some situations where this is not true. For some of these flat fee groups, the initial visit reported occurred in 2018, but the remaining visits that were part of this flat fee group occurred in 2019. In this case, the 2018 flat fee group represented on this file would consist of one event (the stem). The 2019 “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 2017 but subsequent visits occurred during 2018. In this case, the initial visit would not be represented on the file. This 2018 flat fee group would then only consist of one or more leaf records and no stem. Please note that the crosswalk in this document lists all possible flat fee variables.

### **2.5.4 Condition Codes**

Conditions data are not collected for dental events; therefore, this file cannot be linked to the Conditions File.

### **2.5.5 Expenditure Data**

#### **2.5.5.1 Definition of Expenditures**

Expenditures on this file refer to what is paid for dental services. More specifically, expenditures in MEPS are defined as the sum of payments for care received, including out-of-pocket payments and payments made by private insurance, Medicaid, Medicare, and other sources. The definition of expenditures used in MEPS differs slightly from its predecessors, the 1987 NMES and 1977 NMCES surveys, where “charges” rather than sum of payments were used to measure expenditures. This change was adopted because charges became a less appropriate proxy for medical expenditures during the 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. Another general change from the two prior surveys is that charges associated with uncollected liability, bad debt, and charitable care (unless provided by a public clinic or hospital) are not counted as expenditures because there are no payments associated with those classifications. While charge data are provided on this file, data users/analysts should use caution when working with 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 resource costs of those services, nor are they directly comparable to the expenditures defined in the 1987 NMES. For details on expenditure definitions, please reference the following, “Informing American Health Care Policy” (Monheit et al., 1999). AHRQ has developed factors to apply to the 1987 NMES expenditure data to facilitate longitudinal analysis. These factors can be 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 3.3 for more information.

### **2.5.5.2 Data Editing and Imputation Methodologies of Expenditure Variables**

The general methodology used for editing and imputing expenditure data is described below. The MPC did not include either the dental events or other medical expenditures (such as glasses, contact lenses, and hearing devices). Therefore, although the general procedures remain the same for dental and other medical expenditures, editing and imputation methodologies were applied only to household-reported data. Please see below for details on the differences between these editing/imputation methodologies. Separate imputations were performed for flat fee and simple events as well.

#### **2.5.5.2.1 General Data Editing Methodology**

Logical edits were used to resolve internal inconsistencies and other problems in the HC survey-reported data. The edits were designed to preserve partial payment data from households and providers, and to identify actual and potential sources of payment for each household-reported event. In general, these edits accounted for outliers, copayments or charges reported as total payments, and reimbursed amounts that were reported as out-of-pocket payments. In addition, edits were implemented to correct for misclassifications between Medicare and Medicaid, and between Medicare HMOs and private HMOs as payment sources. These edits produced a complete vector of expenditures for some events, and provided the starting point for imputing missing expenditures in the remaining events.

#### **2.5.5.2.2 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 imputations for the flat fee events were carried out separately from the simple events.

The 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 taking into account the persons' weighted distribution in the imputation process.

### **2.5.5.2.3 Dental Data Editing and Imputation**

Expenditures on visits to dentists were developed in a sequence of logical edits and imputations. The household edits were used to correct obvious errors in the reporting of expenditures, and to identify actual and potential sources of payments. Some of the edits were global (i.e., applied to all events); others were hierarchical and mutually exclusive. 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 each covered by a single charge (i.e., simple events). Dental services were imputed as flat fee events if the charges covered a package of health care services (e.g., orthodontia), and all of the services were part of the same event type (i.e., a pure bundle). If a bundle contained more than one type of event, the services were treated as simple events in the imputations (See Section 2.5.3 for more detail on the definition and imputation of events in flat fee bundles.)

Logical edits were also 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 various recipient categories.

The logical edits produced nine recipient categories for events with missing data. Eight of the categories were for events with a common pattern of missing data and a primary payer other than Medicaid. Medicaid events were imputed separately because persons on Medicaid rarely know the provider's charge for services or the amount paid by the state Medicaid program. As a result, the total charge for Medicaid-covered services was imputed and discounted to reflect the amount that a state program would pay for the care.

Separate predictive mean matching imputations were used to impute missing data in each of the eight recipient categories. 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 event among complete events (donors) is not represented among incomplete events (recipients).

### **2.5.5.3 Imputation Flag Variable (IMPFLAG)**

IMPFLAG is a six-category variable that indicates if the event contains complete Household Component (HC) or Medical Provider Component (MPC) data, was fully or partially imputed, or



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

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

IMPFLAG = 1 complete HC data

IMPFLAG = 2 complete MPC data (not applicable to DV events)

IMPFLAG = 3 fully imputed

IMPFLAG = 4 partially imputed

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

#### **2.5.5.4 Flat Fee Expenditures**

The approach used to count expenditures for flat fees was to place the expenditure on the first visit of the flat fee group. The remaining visits have zero payments. Thus, if the first visit in the flat fee group occurred prior to 2018, all of the events that occurred in 2018 will have zero payments. Conversely, if the first event in the flat fee group occurred at the end of 2017, the total expenditure for the entire flat fee group will be on that event, regardless of the number of events it covered after 2018. See Section 2.5.3 for details on the flat fee variables.

#### **2.5.5.5 Zero Expenditures**

As noted above, there are some dental events reported by respondents where the payments were zero. This could occur for several reasons including (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 trial. If all of the medical events for a person fell into one of these categories, then the total annual expenditures for that person would be zero.

#### **2.5.5.6 Sources of Payment**

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

1. Out-of-pocket by User or Family,
2. Medicare,
3. Medicaid,
4. Private Insurance,

5. Veterans Administration/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.

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

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

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

#### **2.5.5.7 Dental Expenditure Variables (DVSF18X - DVTC18X)**

DVSF18X - DVOT18X are the 12 sources of payment. DVXP18X is the sum of the 12 sources of payment for the dental expenditures, and DVTC18X is the total charge. The 12 sources of payment are: self/family (DVSF18X), Medicare (DVMR18X), Medicaid (DVMD18X), private insurance (DVPV18X), Veterans Administration/CHAMPVA (DVVA18X), TRICARE (DVTR18X), other Federal sources (DVOF18X), State and Local (non-federal) government sources (DVSL18X), Workers' Compensation (DVWC18X), other private insurance (DVOR18X), other public insurance (DVOU18X), and other insurance (DVOT18X).

### **2.5.5.8 Rounding**

Expenditure variables on the 2018 dental file have been rounded to the nearest penny. Person-level expenditure information to be released on the MEPS 2018 Full Year Consolidated File will be rounded to the nearest dollar. It should be noted that using the MEPS event files to create person-level totals will yield slightly different totals than those found on the full-year consolidated file. These differences are due to rounding only. Moreover, in some instances, the number of persons having expenditures on the event files for a particular source of payment may differ from the number of persons with expenditures on the person-level expenditure file for that source of payment. This difference is also an artifact of rounding only.

## **3.0 Sample Weight (PERWT18F)**

### **3.1 Overview**

There is a single full year person-level weight (PERWT18F) assigned to each record for each key, in-scope person who responded to MEPS for the full period of time that he or she was in-scope during 2018. A key person was either a member of a responding NHIS household at the time of 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 he or she is a member of the civilian noninstitutionalized portion of the U.S. population.

### **3.2 Details on Person Weight Construction**

The person-level weight PERWT18F was developed in several stages. Person-level weights for Panel 22 and Panel 23 were created separately. The weighting process for each panel included an adjustment for nonresponse over time and calibration to independent population figures. The calibration was initially accomplished separately for each panel by raking the corresponding sample weights for those in-scope at the end of the calendar year to 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 (no degree, high school/GED no college, some college, bachelor's degree or higher); 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. A 2018 composite weight was then formed by multiplying each weight from Panel 22 by the factor .49 and each weight from Panel 23 by the factor .51. The choice of factors reflected the relative sample sizes of the two panels, helping to limit the variance of estimates obtained from pooling the two samples. The composite weight was raked to the same set of CPS-based control totals. When the poverty status information derived from income variables became available, a final raking was undertaken, establishing control figures reflecting poverty status rather than educational attainment. Thus, control totals were established using poverty status (five categories: below poverty, from 100 to 125 percent of poverty, from 125 to 200 percent of poverty, from 200 to 400 percent of poverty, at least 400 percent of poverty) as well as the other five variables previously used in the weight calibration.

### **3.2.1 MEPS Panel 22 Weight Development Process**

The person-level weight for MEPS Panel 22 was developed using the 2017 full year weight for an individual as a “base” weight for survey participants present in 2017. For key, in-scope members who joined an RU some time in 2017 after being out-of-scope in 2017, the initially assigned person-level weight was the corresponding 2017 family weight. The weighting process included an adjustment for person-level nonresponse over Rounds 4 and 5 as well as raking to population control totals for December 2018 for key, responding persons in-scope on December 31, 2018. These control figures were derived by scaling back the population distribution obtained from the March 2019 CPS to reflect the December 31, 2018 estimated population total (estimated based on Census projections for January 1, 2019). Variables used for person-level raking included: educational attainment of the reference person (no degree, high school/GED no college, some college, bachelor’s degree or higher); 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. (Poverty status is not included in this version of the MEPS full year database because of the time required to process the income data collected and then assign persons to a poverty status category). The final weight for key, responding persons who were not in-scope on December 31, 2018 but were in-scope earlier in the year was the person weight after the nonresponse adjustment.

Note that the 2017 full-year weight that was used as the base weight for Panel 22 was derived as follows; adjustment of the MEPS Round 1 weight for nonresponse over the remaining data collection rounds in 2017; and raking the resulting nonresponse adjusted weight to December 2017 population control figures.

### **3.2.2 MEPS Panel 23 Weight Development Process**

The person-level weight for MEPS Panel 23 was developed using the 2018 MEPS Round 1 person-level weight as a “base” weight. For key, in-scope members who joined an RU after Round 1, the Round 1 family weight served as a “base” weight. The weighting process included an adjustment for nonresponse over the remaining data collection rounds in 2018 as well as raking to the same population control figures for December 2018 used for the MEPS Panel 22 weights for key, responding persons in-scope on December 31, 2018. The same six variables employed for Panel 22 raking (educational attainment of the reference person, census region, MSA status, race/ethnicity, sex, and age) were used for Panel 23 raking. Again, the final weight for key, responding persons who were not in-scope on December 31, 2018 but were in-scope earlier in the year was the person weight after the nonresponse adjustment.

Note that the MEPS Round 1 weights for Panel 23 incorporated the following components: the original household probability of selection for the NHIS; proportion of the NHIS sample reserved for MEPS; adjustment for NHIS nonresponse; the probability of selection of NHIS responding households for MEPS; an adjustment for nonresponse at the dwelling unit level for Round 1; and poststratification to U.S. civilian noninstitutionalized population estimates at the family and person level obtained from the corresponding March CPS databases.

### **3.2.3 The Final Weight for 2018**

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, 2018 were adjusted for expected undercoverage. Specifically, the weights of those who were in-scope some time during the year, out-of-scope on December 31, and entered a nursing home during the year and still residing in a nursing home at the end of the year were adjusted to compensate for expected undercoverage for this subpopulation. The weights of persons who died while in-scope during 2018 were poststratified to corresponding estimates derived using data obtained from the Medicare Current Beneficiary Survey (MCBS) and Vital Statistics information provided by the National Center for Health Statistics (NCHS). Separate 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, 2018 is 322,920,490 (PERWT18F>0 and INSC1231=1). The sum of person-level weights across all persons assigned a positive person-level weight is 326,327,888.

### **3.2.4 Coverage**

The target population for MEPS in this file is the 2018 U.S. civilian noninstitutionalized population. However, the MEPS sampled households are a subsample of the NHIS households interviewed in 2016 (Panel 22) and 2017 (Panel 23). New households created after the NHIS interviews for the respective Panels and consisting exclusively of persons who entered the target population after 2016 (Panel 22) or after 2017 (Panel 23) are not covered by MEPS. Neither are previously out-of-scope persons who join an existing household but are unrelated to the current household residents. Persons not covered by a given MEPS panel thus include some members of the following groups: immigrants; persons leaving the military; U.S. citizens returning from residence in another country; and persons leaving institutions. The set of uncovered persons constitutes only a small segment of the MEPS target population.

## **3.3 Using MEPS Data for Trend Analysis**

MEPS began in 1996, and the utility of the survey for analyzing health care trends expands with each additional year of data; however, there are a variety of methodological and statistical considerations when examining trends over time using MEPS. Tests of statistical significance should be conducted to assess the likelihood that observed trends may be attributable to sampling variation. The length of time being analyzed should also be considered. In particular, large shifts in survey estimates over short periods of time (e.g. from one year to the next) that are statistically significant should be interpreted with caution unless they are attributable to known factors such as changes in public policy, economic conditions, or MEPS survey methodology.

With respect to methodological considerations, in 2013 MEPS introduced an effort to obtain more complete information about health care utilization from MEPS respondents with full implementation in 2014. This effort likely resulted in improved data quality and a reduction in underreporting starting in FY 2014 and could have some modest impact on analyses involving trends in utilization across years. The change in the NHIS sample design in 2016 could also potentially affect trend analyses. For example, coverage of the MEPS target population would be

expected to have increased, so subpopulations whose coverage rates were particularly increased would have increased contributions from undercovered portions of their subpopulation.

Another change with the potential to affect trend analyses involved modifications to the MEPS instrument design and data collection process. 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 PUFs were established from data collected in Rounds 1-3 of Panel 22 and Rounds 3-5 of Panel 21, they reflected two different instrument designs. In order 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 were designed to make the data collection effort more efficient and easy 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.

There are also statistical factors to consider in interpreting trend analyses. Looking at changes over longer periods of time can provide a more complete picture of underlying trends. Analysts may wish to consider using techniques to evaluate, smooth, or stabilize analyses of trends using MEPS data such as comparing pooled time periods (e.g. 1996-97 versus 2011-13), working with moving averages, or using modeling techniques with several consecutive years of MEPS data to test the fit of specified patterns over time. Finally, researchers should be aware of the impact of multiple comparisons on Type I error. Without making appropriate allowance for multiple comparisons, undertaking numerous statistical significance tests of trends increases 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 2018 event-level estimates for the U.S. civilian noninstitutionalized population on dental visits as well as expenditures, and sources of payment for these visits. The weight assigned to each dental visit reported is the person-level weight of the person who visited the dentist. If a person reported several visits, each visit is assigned that individual's person-level weight. Estimates of total visits are the sum of the weight variable (PERWT18F) across relevant event records while estimates of other variables must be weighted by PERWT18F to be nationally representative. For example, the appropriate estimate for the mean out-of-pocket payment per dental visit can be represented 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,

$W_j = \text{PERWT18F}_j$  (full year person weight for the person associated with event j), and

$X_j = \text{DVSF18X}_j$  (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 or SPSS. The tables below contain the event-level estimates for several key variables on this file.

### Selected Event-Level Estimates

#### Visits

<b>Estimate of Interest</b>	<b>Variable Name</b>	<b>Estimate (SE)</b>	<b>Estimate Excluding Zero Payment Events (SE)</b>
Total number of dental visits (in millions)	PERWT18F	344.1 (10.91)	291.0 (9.00)
Proportion of dental visits with expenditures > 0*	DVXP18X	0.846 (0.0053)	-----

\*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 trial.

#### Expenditures

<b>Estimate of Interest</b>	<b>Variable Name</b>	<b>Estimate (SE)</b>	<b>Estimate Excluding Zero Payment Events (SE)</b>
Mean total payments per visit	DVXP18X	\$375 (\$7.5)	\$443 (\$9.1)
Mean out-of-pocket payment per visit	DVSF18X	\$171 (\$5.4)	\$202 (\$6.4)
Mean proportion of total expenditures paid by private insurance per visit	DVPV18X/ DVXP18X	-----	0.478 (0.0077)

**Expenditures: Dental Hygienist Visits (DENTHYG = 1)**

<b>Estimate of Interest</b>	<b>Variable Name</b>	<b>Estimate (SE)</b>	<b>Estimate Excluding Zero Payment Events (SE)</b>
Mean total payments per visit where person saw hygienist	DVXP18X	\$242 (\$8.2)	\$252 (\$8.3)
Mean out-of-pocket payment per visit where person saw hygienist	DVSF18X	\$85 (\$5.1)	\$89 (\$5.4)
Mean proportion of total expenditures per visit paid by private insurance where person saw hygienist	DVPV18X/ DVXP18X	-----	0.558 (0.0108)

**4.2 Person-Based Estimates for Dental Care**

To enhance analyses of dental care, analysts may link information about dental visits by sample persons in this file to the annual full year consolidated file (which has data for all MEPS sample persons), or conversely, link person-level information from the full year consolidated file to this event-level file (see Section 5 below for more details). Both this file and the full year consolidated file may be used to derive estimates for persons with dental care and annual estimates of total expenditures. However, if the estimate relates to the entire population, this file cannot be used to calculate the denominator, as only those persons with at least one dental visit are represented on this data file. Therefore, the full year consolidated file must be used for person-level analyses that include both persons with and without dental 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 estimated, 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 include or exclude such cases in the numerator and/or denominator when calculating proportions.

Methodologies used for the editing/imputation of expenditure variables (e.g., sources of payment, flat fee, and zero expenditures) are described in Section 2.5.5.2.

**4.4 Variance Estimation (VARPSU, VARSTR)**

MEPS has a complex sample design. To obtain estimates of variability (such as the standard error of sample estimates or corresponding confidence intervals) for MEPS estimates, analysts need to take into account the complex sample design of 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, and jackknife replication. Various software packages provide analysts with the capability of implementing these methodologies. MEPS analysts most commonly use the Taylor Series approach. However, the capability of employing the Balanced Repeated Replication (BRR) methodology is also provided if needed to develop variances for more complex estimators.

#### **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 as well as all other MEPS public use files. Software packages that permit the use of the Taylor-series linearization method include SUDAAN, Stata, SAS (version 8.2 and higher), and SPSS (version 12.0 and higher). For complete information on the capabilities of each package, analysts should refer to the corresponding software user documentation.

Using 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 MEPS data file serve to 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 computer software packages will provide estimated standard errors appropriate for assessing the variability of MEPS survey estimates. It should be noted that the number of degrees of freedom associated with estimates of variability indicated by such a package may not appropriately reflect the number available. For variables of interest distributed throughout the country (and thus the MEPS sample PSUs), one can generally expect to have at least 100 degrees of freedom associated with the estimated standard errors for national estimates based on this MEPS database.

Prior to 2002, MEPS variance strata and PSUs were developed independently from year to year, and the last two characters of the strata and PSU variable names denoted the year. However, beginning with the 2002 Point-in-Time PUF, the variance strata and PSUs were developed to be compatible with all future PUFs until the NHIS design changed. Thus, when pooling data across years 2002 through the Panel 11 component of the 2007 files, the variance strata and PSU variables provided can be used without modification for variance estimation purposes for estimates covering multiple years of data. There were 203 variance estimation strata, each stratum with either two or three variance estimation PSUs.

From Panel 12 of the 2007 files, a new set of variance strata and PSUs were 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, starting from Panel 12. Therefore, there are a total of 368 (203+165) variance strata in the 2007 Full Year file 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 are based on the next NHIS design, there are 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 in the event that data are pooled for several years.

As discussed, a complete change was made to the NHIS sample design in 2016, effectively changing the MEPS design beginning with calendar year 2017. Both Panels 22 and 23 reflect this new design. There were 117 variance strata originally formed under this new design intended for use until the next fully new NHIS design was implemented. They appear in the various MEPS data sets associated with 2017 as well as for the 2018 Point-in-Time PUF involving both Panels 22 and 23. However, it was later learned that the NHIS sample design was further modified in 2018, calling for a reconstruction of the previously established variance strata. Technically, this reconstruction would not be required until the MEPS 2019 PUFs were to be constructed. However, some analysts pool MEPS data across several years. In order to accommodate such pooling, the modification to the MEPS variance structure is being implemented initially for this 2018 FY PUF. Only a handful of variance strata have been affected with some pooling of previous strata being necessary. There are now 110 variance strata established for MEPS, compared to the 117 previously established.

In order to make the pooling of data across multiple years of MEPS more straightforward, the numbering system for the variance strata has changed. Those strata associated with the new design have four digit values with a “2” as the first digit. Those associated with the previous design have “1” as the first of four digits.

If analyses call for pooling MEPS data across several years, in order to ensure that variance strata are identified appropriately for variance estimation purposes, one can proceed as follows:

1. When pooling any year from 2002 or later, one can use the variance strata numbering as is.
2. When pooling any year from 1996 to 2001 with any year from 2002 or later, use the H36 file.
3. A new H36 file was constructed to allow pooling of 2007 and later years with 1996 to 2006.

#### **4.4.2 Balanced Repeated Replication (BRR) 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 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 to compute variances of complex non-linear estimators for which a Taylor linear form is not easy to derive and not available in commonly used software. For instance, it is not possible to calculate the variances of a median or the ratio of two medians using the Taylor linearization method. For these types of estimators, users may 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 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 data set, HC-036BRR, contains the information necessary to construct the BRR replicates. It contains 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 analysis of MEPS data pooled across years, the BRR replicates can be formed in the same way using the HC-036 file. For more information about creating BRR replicates, users can refer to the documentation for the HC-036BRR pooled linkage file.

## **5.0 Merging/Linking MEPS Data Files**

Data from this file can be used alone or in conjunction with other files for different analytic purposes. This section summarizes various scenarios for merging/linking MEPS event files. The set of households selected for MEPS is a subsample of those participating in the National Health Interview Survey (NHIS), thus, each MEPS panel can also be linked back to the previous year's NHIS public use data files. For information on obtaining MEPS/NHIS link files please see the [MEPS website](#).

### **5.1 Linking to the Person-Level File**

Merging characteristics of interest from other MEPS files (e.g., 2018 Full Year Consolidated File or 2018 Prescribed Medicines File) expands the scope of potential estimates. For example, to estimate the total number of dental events of persons with specific demographic characteristics (such as age, race, and sex), population characteristics from a person-level file need to be merged onto the Dental file. This procedure is shown below. The MEPS 2018 Appendix File, HC-206I, provides additional details of how to merge other MEPS data files.

1. Create data set PERSX by sorting the 2018 Full Year Consolidated File, by the person identifier, DUPERSID. Keep only variables to be merged onto the Dental file and DUPERSID.
2. Create data set DENT by sorting the dental event file by person identifier, DUPERSID.
3. Create final data set NEWDENT by merging these two files by DUPERSID, keeping only records on the dental event file.

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

```
PROC SORT DATA=HCXXX (KEEP=DUPERSID AGE31X AGE42X AGE53X  
SEX RACEV1X EDUCYR HIDEG) OUT=PERSX;  
    BY DUPERSID;  
RUN;  
  
PROC SORT DATA=DENT;  
    BY DUPERSID;
```

```
RUN;  
  
DATA NEWDENT;  
    MERGE DENT (IN=A) PERSX (IN=B);  
    BY DUPERSID;  
IF A;  
RUN;
```

The MEPS 2018 Appendix File, HC-206I, provides examples of how to merge other MEPS data files.

## **5.2 Linking to the Prescribed Medicines File**

The RXLK file provides a link from the MEPS event files to the 2018 Prescribed Medicines Event File. When using RXLK, data users/analysts should keep in mind that one dental visit can link to more than one prescribed medicine record. Conversely, a prescribed medicine event may link to more than one dental visit or different types of events. When this occurs, it is up to the data user/analyst to determine how the prescribed medicine expenditures should be allocated among those medical events. For detailed linking examples, including SAS code, data users/analysts should refer to the MEPS 2018 Appendix File, HC-206I.

## **5.3 Linking to the Medical Conditions File**

Conditions data are not collected for dental events; therefore, this file cannot be linked to the Conditions File.

## References

- Chowdhury, S.R., Machlin, S.R., Gwet, K.L. [Sample Designs of the Medical Expenditure Panel Survey Household Component, 1996–2006 and 2007–2016. Methodology Report #33.](#) January 2019. Agency for Healthcare Research and Quality, Rockville, MD.
- Cohen, S.B. (1996). The Redesign of the Medical Expenditure Panel Survey: A Component of the DHHS Survey Integration Plan. *Proceedings of the COPAFS Seminar on Statistical Methodology in the Public Service.*
- Fay, R.E. (1989). Theory and Application of Replicate Weighting for Variance Calculations. *Proceedings of the Survey Research Methods Sections, ASA*, 212-217.
- Monheit, A.C., Wilson, R., and Arnett, III, R.H. (Editors) (1999). *Informing American Health Care Policy.* Jossey-Bass Inc, San Francisco.
- Shah, B.V., Barnwell, B.G., Bieler, G.S., Boyle, K.E., Folsom, R.E., Lavange, L., Wheelless, S.C., and Williams, R. (1996). *Technical Manual: Statistical Methods and Algorithms Used in SUDAAN Release 7.0.* Research Triangle Park, NC: Research Triangle Institute.

## **D. Variable-Source Crosswalk**

**VARIABLE-SOURCE CROSSWALK**

**FOR MEPS HC-206B: 2018 DENTAL VISITS**

**Survey Administration Variables**

<b>Variable</b>	<b>Description</b>	<b>Source</b>
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
FFEEIDX	Flat fee ID	CAPI derived
PANEL	Panel Number	Constructed

**Dental Events Variables**

<b>Variable</b>	<b>Description</b>	<b>Source</b>
DVDATEYR	Event date – year	CAPI derived
DVDATEMM	Event date – month	CAPI derived
GENDENT_M18	General dentist seen	DN10
DENTHYG_M18	Dental hygienist seen	DN10
DNSPCLST_M18	Dental specialist seen	DN10
PEDDENT_M18	Pediatric dentist seen	DN10
DENTYPE_M18	Other dental specialist seen	DN10
EXAMINE_M18	General exam, checkup or consultation	DN20
CLENTETX	Edited CLENTETH_M18	DN20 (Edited)
CLENTETH_M18	Cleaning, prophylaxis, polishing or periodontal recall	DN20
JUSTXRYX	Edited JUSTXRAY_M18	DN20 (Edited)
JUSTXRAY_M18	X-rays, radiographs or bitewings	DN20
FLUORIDE_M18	Fluoride treatment	DN20
SEALANTX	Edited SEALANT_M18	DN20 (Edited)
SEALANT_M18	Sealant application	DN20
FILLINGX	Edited FILLING_M18	DN20 (Edited)
FILLING_M18	Fillings, inlays, crowns or caps	DN20

### Dental Events Variables

<b>Variable</b>	<b>Description</b>	<b>Source</b>
ROOTCANX	Edited ROOTCANL_M18	DN20 (Edited)
ROOTCANL_M18	Root canal	DN20
GUMSURGX	Edited GUMSURG_M18	DN20 (Edited)
GUMSURG_M18	Periodontal scaling, root planing or gum surgery	DN20
IMPLANTX	Edited IMPLANT_M18	DN20 (Edited)
IMPLANT_M18	Implants	DN20
ORALSURX	Edited ORALSURG_M18	DN20 (Edited)
ORALSURG_M18	Extraction, tooth pulled or oth oral surgery	DN20
BRIDGESX	Edited BRIDGES_M18	DN20 (Edited)
BRIDGES_M18	Fixed or relining/repair of bridges/dentures, removable dentures	DN20
ORTHDONX	Edited ORTHDONT_M18	DN20 (Edited)
ORTHDONT_M18	Orthodontia, braces or retainers	DN20
DENTPROX	Edited DENTPROC	DN20OS (Edited)
DENTPROC	Other dental procedures	DN20OS
DENTOTHX	Other specify dental procedures edited	DN20OS (Edited)

### Flat Fee Variables

<b>Variable</b>	<b>Description</b>	<b>Source</b>
FFDVTYPE	Flat fee bundle	Constructed
FFBEF18	Total # of visits in FF before 2018	FF50
FFTOT19	Total # of visits in FF after 2018	FF60

### Imputed Expenditure Variables

<b>Variable</b>	<b>Description</b>	<b>Source</b>
DVSF18X	Amount paid, family (Imputed)	CP Section (Edited)
DVMR18X	Amount paid, Medicare (Imputed)	CP Section (Edited)
DVMD18X	Amount paid, Medicaid (Imputed)	CP Section (Edited)
DVPV18X	Amount paid, private insurance (Imputed)	CP Section (Edited)
DVVA18X	Amount paid, Veterans/CHAMPVA (Imputed)	CP Section (Edited)
DVTR18X	Amount paid, TRICARE (Imputed)	CP Section (Edited)



### Imputed Expenditure Variables

<b>Variable</b>	<b>Description</b>	<b>Source</b>
DVOF18X	Amount paid, other federal (Imputed)	CP Section (Edited)
DVSL18X	Amount paid, state & local government (Imputed)	CP Section (Edited)
DVWC18X	Amount paid, workers' comp (Imputed)	CP Section (Edited)
DVOR18X	Amount paid, other private (Imputed)	Constructed
DVOU18X	Amount paid, other public (Imputed)	Constructed
DVOT18X	Amount paid, other insurance (Imputed)	CP Section (Edited)
DVXP18X	Sum of DVSF18X – DVOT18X (Imputed)	Constructed
DVTC18X	Household reported total charge (Imputed)	CP Section (Edited)
IMPFLAG	Imputation status	Constructed

### Weights Variables

<b>Variable</b>	<b>Description</b>	<b>Source</b>
PERWT18F	Expenditure File Person Weight, 2018	Constructed
VARSTR	Variance estimation stratum, 2018	Constructed
VARPSU	Variance estimation PSU, 2018	Constructed