METHODOLOGY REPORT #36

February 24, 2023

Analyzing Medical Conditions in MEPS: Detailed Reference Guide (Part 2 of 2)

Emily Mitchell, PhD, Rebecca Ahrnsbrak, MPS, Anita Soni, PhD, and Steve Machlin, MS

Summary

This report is the second part of a two-part series of documents designed to help users analyze medical conditions in the Medical Expenditure Panel Survey (MEPS) and understand changes to the survey that may affect analyses. Over time, MEPS has focused increasingly on asking respondents about conditions associated with medical care and prescription drugs, as well as chronic medical conditions deemed a priority for research due to their high prevalence, rather than collecting data on all possible conditions. This detailed reference guide documents changes to the survey that have affected how data on medical conditions are collected, processed, and provided to the public in the form of public-use files. Part 1 of this series is a short user guide that provides specific recommendations for analyzing MEPS condition data.

Suggested Citation

Mitchell, E., Ahrnsbrak, R., Soni, A., and Machlin, S. *Analyzing Medical Conditions in MEPS: Detailed Reference Guide (Part 2 of 2).* Methodology Report #36. Rockville, MD: Agency for Healthcare Research and Quality; February 2023.
http://www.meps.ahrq.gov/mepsweb/data%20files/publications/mr36/mr36_dr.pdf

Acknowledgments

The authors would like to thank Steve Hill for his careful review and thoughtful suggestions; Zhengyi Fang for programming support and quality control; Jerrod Anderson for his review and input; and Jeannie Hudnall, Angie Kistler, and Joanne Campione for their expertise and careful review of the technical details of MEPS survey administration and data processing.

* * *

Center for Financing, Access, and Cost Trends
Agency for Healthcare Research and Quality
5600 Fishers Lane, Mailstop 07W41A
Rockville, MD 20857
http://www.meps.ahrq.gov

Table of Contents

Analyzing Medical Conditions in MEPS: Detailed Reference Guide (Part 2 of 2)

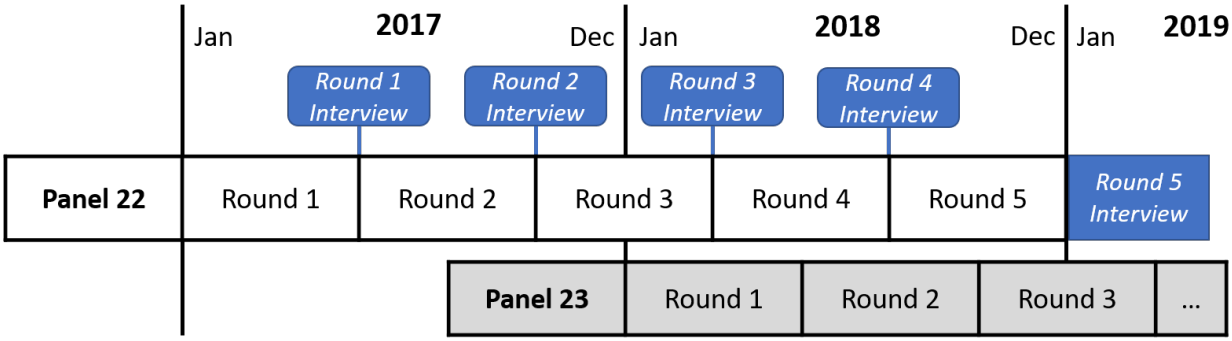
Emily Mitchell, PhD, Rebecca Ahrnsbrak, MPS, Anita Soni, PhD, and Steve Machlin, MS

Background

The Medical Expenditure Panel Survey (MEPS) Household Component (HC) provides national estimates of healthcare expenditures and utilization for a nationally representative sample of the U.S. civilian noninstitutionalized population. The MEPS Household Component (MEPS-HC) includes information on sources of payment, demographics, health insurance coverage, and medical conditions. The National Health Interview Survey (NHIS) serves as the sampling frame for the MEPS-HC. The NHIS is conducted by the National Center for Health Statistics (Cohen, 1997). The MEPS-HC collects data from a nationally representative sample of households through an overlapping panel design. A new panel of sample households is selected each year, and data for each panel are collected for two calendar years (Ezzati-Rice, Rohde, and Greenblatt, 2008). Data for each panel are collected through a series of five rounds of interviews that take place over a 2.5-year period.[1] This approach provides continuous and current estimates of healthcare expenditures at both the person and household level for each calendar year.

Figure 1 illustrates the typical timing and relationship between panels, rounds of interviews, and calendar years. For example, looking at data collection by panel, Rounds 3–5 of Panel 22 and Rounds 1–3 of Panel 23 supply the data for 2018.

Figure 1. Timing of Panels, Rounds, and Interviews



Data collection for the MEPS-HC includes many types of survey questions, some of which only pertain to subsets of the diverse respondents participating in the survey. To accommodate the extensive array of questions asked, data are collected using an intricate system of skip patterns and questionnaire modules grouped into sections. Computer-assisted personal interviewing (CAPI) using a laptop computer makes it possible to field such a complex data collection instrument. The CAPI data collection instrument consists of sections that are composed of a series of computer screens, which contain questions and interviewing instructions that can vary across rounds. Some sections are included in every round of data collection, while other sections are only included in one or two rounds. Once data collection and editing are completed, the MEPS data are released to the public in the form of public-use data files, tables, and

[1] Starting with data year 2020, additional rounds of data collection were added to the typical five rounds of interviews to counteract the lower response rates due to the COVID-19 pandemic. As a result, full-year files for 2020 contain data from three panels, full-year files for 2021 contain data from four panels, and full-year files for 2022 contain data from three panels. Refer to Documentation for the 2020 Full-Year Consolidated file for more information.

interactive data tools via the MEPS website (meps.ahrq.gov) and AHRQ data tools website (datatools.ahrq.gov/meps-hc).

Over time, MEPS has focused increasingly on asking respondents about conditions that are associated with medical care and prescription drugs (i.e., *treated* conditions), and chronic conditions that AHRQ has designated as a priority to research due to their high prevalence, rather than casting a wider net. These changes reduced respondent burden by removing questions about rare conditions in the general population (for example, dental injuries) and conditions that were not treated and did not benchmark to external data sources. This report describes changes in how household respondents are asked about medical conditions and how their responses are coded. This document focuses on the changes that could make a notable impact on trend analyses for conditions that use the Medical Condition files.

Condition Data Collection and Public-Use Files

Each year, MEPS releases public-use files (PUFs) after data collection and processing are completed. Two of these files, the Full-Year Consolidated file and the Medical Conditions file, contain data on health status and medical conditions as reported by MEPS respondents.

The **Full-Year Consolidated PUF** is a person-level file that contains data on demographics, healthcare spending, health status, and priority conditions. The health status questions ask whether perceived health and mental health are considered excellent, very good, good, fair, or poor. Priority conditions are a set of 15 medical conditions that tend to be chronic in nature and which AHRQ has deemed a priority for research due to their high prevalence. Most priority condition questions ask respondents whether household members were *ever* diagnosed with a particular condition, while other questions ask about diagnoses in the past year. Follow-up questions such as age of diagnosis are also asked where applicable. For most priority conditions, questions are only asked of adults ages 18 and older; however, for diabetes and asthma, questions are asked for respondents of all ages. In addition, for attention-deficit/hyperactivity disorder (ADHD), questions are asked for children ages 5–17. A complete list of person-level health and condition variables in the Full-Year Consolidated PUFs can be found in Appendix A.

The **Medical Conditions PUF** is a condition-level file, which contains a record for each reported medical condition for each household member. These files include additional information about each medical condition, including classification codes from the *International Classification of Diseases, Tenth Revision, Clinical Modification* (ICD-10-CM) and the Clinical Classification Software Refined (CCSR).[2] For the data years 1996–2015, previous versions of ICD and Clinical Classification Software (CCS) codes were released. Condition records can be linked to the Full-Year Consolidated file and to the MEPS event files.

A complete list of the variables in the Medical Condition PUF can be found in Appendix B. There are several nuances of the Medical Conditions files that users should keep in mind:

[2] ICD-10-CM: *International Classification of Diseases*, *Tenth Revision, Clinical Modification* (https://cdc.gov/nchs/icd/icd-10-cm.htm). CCSR: Clinical Classification Software Refined (https://hcup-us.ahrq.gov/toolssoftware/ccsr/ccs_refined.jsp).

- Some conditions may appear to be "duplicate" conditions at the ICD level. This could occur, for instance, if the fully specified ICD-9 or ICD-10 code differs but the first three digits are the same.[3]

- The **NUM variables in the Medical Conditions PUF (HHNUM, IPNUM, OPNUM, OBNUM, ERNUM, and RXNUM) were constructed by counting records in the event files. These variables can be used to identify which event types are associated with a particular condition but they should not be used for utilization estimates. For instance, the HHNUM variable includes informal healthcare counts, OBNUM and OPNUM include phone calls (for 1996–2017 data), and RXNUM represents a count of distinct prescribed medicines for a person in a given round on the PMED event file, not the number of fills or the number of drugs.

- Researchers are advised to use caution when comparing MEPS estimates to estimates from other surveys and sources. MEPS estimates might not benchmark to other sources due to several reasons:

    - **Survey population:** MEPS data only represent the U.S. civilian noninstitutionalized population. This means, for example, that military personnel and people in nursing homes are not included in MEPS estimates.

    - **Recall error:** MEPS conditions are reported by household respondents and as such, are subject to recall error. Respondents are encouraged to keep records of their medical visits and prescribed medicine purchases, but some respondents are more diligent about recordkeeping than others. In addition, respondents tend to accurately report conditions that are highly salient, cause pain, need ongoing treatment or alter lifestyle, and/or affect daily life (e.g., a heart attack, diabetes, joint pain, etc.). Less salient conditions (e.g., the common cold, the flu) may be underreported due to recall error (Machlin et al., 2009).

    - **Proxy-reported conditions:** MEPS typically has one respondent per household; this respondent responds for every member of the household. The respondent may not have complete information for all people in the household. This can be particularly problematic in households with many adult members, who may not communicate complete information about their medical conditions, visits, and costs to the respondent.

    - **Condition accuracy:** Conditions reported in MEPS often lack the specificity and level of detail that would be commonly found in patient records from a doctor's office or in administrative data. Data about reported conditions may be vague or ambiguous. This may be intentional (e.g., because a respondent may not feel comfortable sharing details about a sensitive condition) or unintentional (e.g., because the respondent may lack detailed information). Prior to 2020, interviewers recorded reported conditions as verbatim text strings, which would later be mapped to ICD-9 or ICD-10 codes. The addition of the conditions pick-list tool in 2020 (see below for more details) helped alleviate some of these issues by providing a searchable list of conditions that the interviewers could choose from. The list also allowed respondents to help clarify

[3] For confidentiality purposes, only the first three digits of the ICD-9 or ICD-10 codes are released in the Medical Conditions PUF.

ambiguous responses in real time (e.g., does "flu" mean "stomach flu" or "influenza"?). Despite these improvements, the accuracy and specificity of reported conditions in MEPS remains limited by what the respondents know and what they choose to report.

Respondents report conditions in several sections of the MEPS questionnaire. When a respondent reports a medical condition, that condition is added to a person's dynamic conditions roster. The conditions roster serves to facilitate subsequent survey sections by prepopulating relevant questions with the reported conditions that can be selected (e.g., as the reason for a visit). Whenever questions use the conditions roster, respondents can report new conditions, which interviewers add to the roster. Conditions in a person's roster are then included in the Medical Conditions PUF if they are *current* conditions (i.e., linked to a medical visit or prescribed medicine during the survey year). Although the survey sections that generate condition records have changed over the years, starting with data year 2018, conditions can be added to the MEPS conditions roster in three ways:

1. A condition can be reported in the **Priority Condition Enumeration (PE)** section. In this section, people are asked if they have ever been diagnosed with specific conditions (Note: for joint pain or chronic bronchitis, people are asked if they have been diagnosed in the last 12 months). These questions are included as "yes" or "no" questions in the Full-Year Consolidated PUF. In addition, a condition reported in the PE section will be added to the conditions roster, which will in turn populate questions asking about healthcare utilization in the Medical Events (ME) and Prescribed Medicines (PM) sections. Conditions reported in the PE section will be included in the Medical Conditions PUF only if they are also linked to medical events or prescribed medicine purchases.

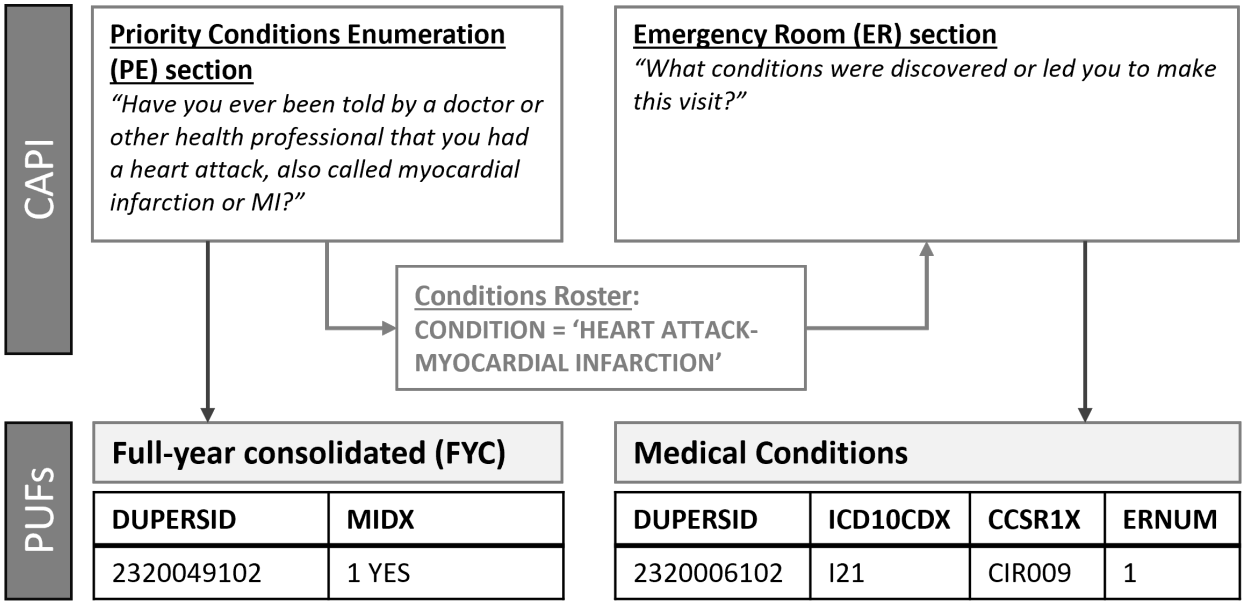
2. A condition can be reported by the household respondent in one of the **ME sections** as the reason for a particular medical event or prescribed medicine purchase. The ME sections are Emergency Room (ER), Home Health (HH), Hospital Stay (HS), Medical Provider Visits (MV), Outpatient Department (OP), Prescribed Medicines (PM), and Telehealth (TH; as of 2020). Other CAPI sections that collect information on medical events and expenditures include Dental Care (DN), Institutional Care Stay (IC), and Other Medical Expenses (OM). The OM section does not collect information on medical conditions. In addition, condition questions were dropped from the DN section in 2002, and conditions from the IC section are not included in the PUFs unless associated with a different event type or prescription medicine.

3. Instead of being reported as a response to a specific question in the survey, a condition could be generated based on a "comment" added by the interviewer during the survey. Historically, this might occur when a respondent mentions a condition during an unrelated portion of the interview. If the interviewer is unable to return to the relevant sections, they could enter a comment about the condition, which would then be reviewed during data processing and editing. This method of adding conditions has become much less common after the recent CAPI design improvements, which have made it easier for interviewers to jump between survey sections to make corrections or add new information as the interview progresses.

The following schematic (Figure 2) illustrates the questionnaire flow for the priority condition "Heart Attack/Myocardial Infarction (MI)." In this example based on 2019 data, the respondent answered "yes" in the PE section when asked if a health professional had ever told them they had a heart attack; their response was recorded in the Full-Year Consolidated PUF (MIDX = "1 YES"). "HEART ATTACK–

MYOCARDIAL INFARCTION" is then added to their conditions roster, which in turn prepopulates subsequent sections of the survey. If the respondent reports that a medical visit or prescribed medicine purchase was associated with their heart attack, then a condition record is created in the Medical Conditions PUF. In this example, the respondent reported that their heart attack was associated with an ER visit (ERNUM = 1), so this condition is also included in the Medical Conditions PUF (ICD10CDX = I21, CCSR1X = CIR009).

Figure 2. Example Schematic From 2019 Data Demonstrating Relationship Between Survey Sections, Conditions Roster, and PUFs



Although the PE and ME sections are the only questionnaire sections that generate conditions in the conditions roster (as of 2018),[4] additional sections have been used to collect data on medical conditions in previous years. The following list describes each survey section that has historically created medical conditions in the medical conditions roster. As the survey has evolved over time, some of these sections have been omitted, while others have been edited to no longer collect condition information. The sections are listed in the order in which they were asked during the interview. The year ranges associated with each of these sections represent the years in which each survey section collected information on medical conditions:

- **Condition Enumeration (CE), 1996–2017:** This section identified specific physical and mental health conditions from verbatim text responses to the question "Between {START DATE} and {END DATE}, did {PERSON} have any physical or mental health problems, accidents, or injuries? What did {PERSON} have?" A household member did not necessarily need to seek treatment or take medicine to include the condition in this section. Starting in 2018, this section

[4] Condition information is also collected in a follow-back survey of medical providers called the Medical Provider Component (MPC) of MEPS. The MPC survey is conducted to obtain more accurate expenditure estimates for healthcare events and is fielded for a sample of medical providers reported by household interview respondents (Stagnitti et al., 2018). Because the condition information obtained from the MPC is not available for everyone in the sample, it is not used to supplement, replace, or verify household-reported condition data.

was no longer fielded because the conditions were not systematically reported and did not benchmark to estimates from other sources.

- **Pregnancy detail (PG), 1996–2007:** This section collected additional information for women identified in the CE section as having been pregnant at any time during the reference period. Starting in 2008, this section was no longer fielded.

- **Dental care (DN): 1996–2001:** Beginning in 2002, questions related to dental injuries (and related medical conditions) were omitted from the DN section.

- **Long-term care (LC), 1997–1998:** This is a supplemental section that collected detailed information on disabilities and functional limitations including related medical conditions. This section was only fielded in 1997 and 1998.

- **Alternative/Preventive Care (AP), 1996–1998:** From 1996 to 1998, this section gathered information on alternative and preventive care received, such as acupuncture, massage therapy, and homeopathic treatment. This section included the question "For what health problems was the alternative care practitioner consulted?" This section was not fielded in 1999. In 2000, alternative care was eliminated from this section, and the condition question was dropped. Questions about alternative care were shifted to the ME sections.

- **Disability Days (DD), 1996–2012:** This section asked about conditions that caused the respondent to miss at least half a day of work or school, or to spend at least half a day in bed. Starting in 2013, the MEPS no longer asked about specific medical conditions in this section.

In earlier years of the MEPS, additional survey sections collected information on medical conditions, but the sections were not connected to the conditions roster. These include the Over-the Counter Medicine (OC) section, the Priority Conditions Quality Supplement (PC), and the Accidents/Injuries and Conditions (CN) section. The OC section was fielded from 1996 to 2001 and collected details about health conditions related to purchases of over-the-counter medicines. The PC section was conducted from 2000 to 2017 and asked about priority medical conditions but did not add them to the conditions roster. The CN section was fielded from 1996 to 2017 and collected additional information on health conditions identified in the ME or DD sections.

Additional information on each survey section of the MEPS-HC can be found at the following links:

- Summary of HC Questionnaire Sections: https://meps.ahrq.gov/mepsweb/survey_comp/hc_ques_sections.jsp
- MEPS Questionnaire Sections: https://meps.ahrq.gov/mepsweb/survey_comp/survey.jsp#Questionnaires

Table 1 shows the weighted number of records in the Medical Conditions files for 1996–2020 based on the survey section that generated the condition (i.e., added it to a person's conditions roster). This table includes weighted counts of the total number of records in the Medical Conditions file, as well as counts limited to *treated* conditions, where treated conditions are conditions that were discovered or led to a medical event or prescribed medicine purchase.

Table 1. Weighted Number of Condition Records (in Millions) in the MEPS Annual Medical Conditions Files First Reported in Each Questionnaire Section, 1996–2020

Year	PE	CE	PG	ME[1]	DN	TH	LC	AP	DD	Other[2]	Total Records	Total Treated[3]
1996	--	510.6	7.2	344.8	2.8	--	--	2.6	85.8	7.5	961.2	647.8
1997	--	455.5	6.8	368.0	2.6	--	3.2	1.7	58.3	7.1	903.3	622.0
1998	--	411.4	6.7	354.4	2.6	--	3.4	2.3	61.0	6.6	848.3	625.3
1999	--	398.4	6.4	360.9	2.4	--	0.6*	0.9*	62.2	6.6	838.4	632.7
2000	--	384.3	7.6	389.9	2.8	--	--	--	61.2	8.2	854.0	646.0
2001	--	403.9	7.3	438.1	2.6	--	--	--	64.8	8.4	925.0	703.5
2002	--	416.9	6.9	472.4	0.1*	--	--	--	62.4	7.4	966.3	745.2
2003	--	428.1	7.5	484.7	--	--	--	--	64.9	8.2	993.4	762.2
2004	--	428.0	7.7	489.9	--	--	--	--	65.2	8.7	999.6	766.0
2005	--	432.3	7.5	493.6	--	--	--	--	65.3	10.8	1,009.5	777.0
2006	--	424.8	7.6	483.9	--	--	--	--	67.4	11.6	995.3	764.0
2007	106.3*	359.8	4.1	450.4	--	--	--	--	60.4	13.0	994.1	776.1
2008	252.6	315.4	--	430.6	--	--	--	--	77.4	13.1	1,089.2	804.0
2009	251.3	327.8	--	454.2	--	--	--	--	75.0	12.3	1,120.6	824.4
2010	255.5	312.2	--	468.9	--	--	--	--	74.6	12.7	1,124.0	835.6
2011	263.0	326.3	--	480.0	--	--	--	--	72.8	12.6	1,154.7	863.0
2012	265.8	349.3	--	489.3	--	--	--	--	74.6	12.0	1,191.0	875.0
2013	265.8	395.5	--	532.1	--	--	--	--	63.4*	14.6	1,271.3	931.4
2014	274.5	419.3	--	552.9	--	--	--	--	5.1*	14.7	1,266.4	964.4
2015	281.5	441.0	--	568.1	--	--	--	--	--	13.9	1,304.4	989.9
2016	278.4	451.3	--	565.4	--	--	--	--	--	13.3	1,308.4	976.3
2017	254.9	354.1*	--	603.6	--	--	--	--	--	9.4	1,222.0	955.5
2018	216.1	42.9*	--	716.7	--	--	--	--	--	1.5	977.2	977.2
2019	210.9	--	--	746.5	--	--	--	--	--	0.02	957.4	957.4
2020	208.6	--	--	703.2	--	8.2	--	--	--	0.02	920.1	920.1

Note. AP = Alternative/Preventive Care; CE = Condition Enumeration; DD = Disability Days; DN = Dental Care; LC = Long-term Care; ME = Medical Events; PE = Priority Condition Enumeration; PG = Pregnancy Detail; TH = telehealth.

[1] The Medical Events (ME) column consists of the following survey sections: Emergency Room (ER), Hospital Stay (HS), Institutional Care Stay (IC), Medical Provider Visits (MV), Outpatient Department (OP), and Prescribed Medicines (PM).

[2] From 1996 to 2017, the vast majority of conditions reported in the "Other" category include conditions derived from comments rather than direct responses to survey questions. Since the 2018 redesign, this method of reporting conditions has become less common, and the remaining conditions in this category are added during analytic editing.

[3] Total treated: Total number of condition records that link to a record in at least one of the event files for that year.

* Transition years of condition collection, where the specified section was asked for only one panel or persisted from the prior year of data collection.

Because the MEPS focuses on healthcare service use and expenditures, the best use of the Medical Conditions file is to analyze *treated* conditions, particularly when comparing trends across years. For the tables in this document, treated conditions are defined as conditions in the Medical Conditions file that link to a record in one of the event files. Note that this includes events with inapplicable expenditures (e.g., outpatient or office-based phone calls or informal home health care). In practice, some analysts may wish to exclude these types of events when identifying treated conditions.

Even though conditions can be reported in more than one section (e.g., the CE and ME sections), the columns in Table 1 are mutually exclusive because they indicate the section in which the condition was *first* reported. Columns are sorted from left to right in the order that each section is fielded. For example, questions in the Priority Conditions Enumeration (PE) section are asked before those in the Condition Enumeration (CE) section. The exception is the "Other" section, which mainly consists of conditions derived from comments and which can be added at multiple points during the interview.

Because of the nature of the overlapping panel design used in MEPS, conditions that were reported in a particular section will persist in the conditions roster throughout additional years of the survey, even if those survey sections are no longer fielded. For instance, if a Panel 6 (2001–2002) respondent reported a condition in the Dental Care (DN) section in Round 1 (2001), then that condition would persist in the person's conditions roster for subsequent rounds and could therefore be included in the Medical Conditions PUF in 2002, even though the Dental Care section no longer collected information on specific conditions in 2002.

Timeline of Changes Affecting the Medical Conditions PUFs

As described above, changes to the MEPS over time have affected the way that conditions are reported in the survey. Many of these changes stem from revisions to survey questions asking about conditions. During the last few years, in addition to the survey section changes, some changes have been made to data collection and processing for conditions, which have also affected the MEPS Medical Condition data. These include:

- Moving from ICD-9/CCS to ICD-10/CCSR codes in 2016
- Adding the conditions pick-list tool and automated probes starting in 2020

Figure 3 illustrates the timing of changes to MEPS that have affected the Medical Conditions PUFs over time. Table 2 summarizes these changes and their impact on data in the PUFs.

Figure 3. Timeline of MEPS Changes That Have Affected the Collection and Processing of Data for Medical Conditions

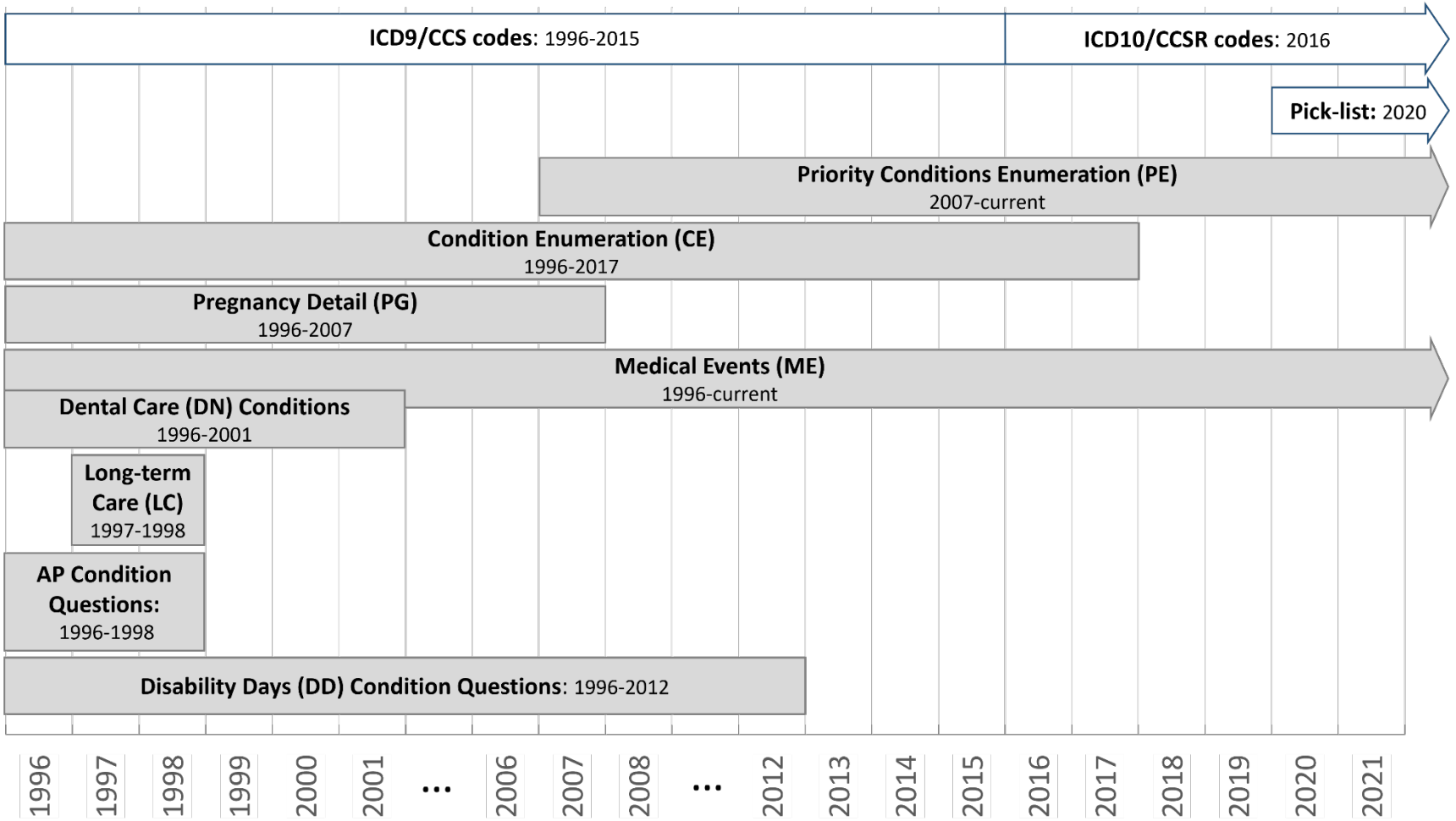


Table 2. Summary of Changes Affecting the MEPS Medical Conditions PUFs

Year	Change	Impact on Reported Conditions	Variables Dropped or Replaced
2020	Conditions pick-list tool and automated probes	Little to no impact on reported conditions; small reduction in "uncodeable" conditions.	None
2018	CE section dropped	Noticeable drop in number of records in Medical Conditions PUF, particularly the elimination of conditions that were reported as "bothering" the sample member but for which medical care was not obtained. Analyses should be limited to *treated* conditions.	None
2016	New Condition Classification coding (ICD-9 to ICD-10)	Caution is needed when comparing conditions from 1996 to 2015 with those from 2016 and later. Apparent discontinuities may be an artifact of the coding transition rather than a true difference.	**ICD9CODX, ICD9PROX,** and **CCCODEX** replaced with **ICD10CDX, CCSR1X-CCSR3X** on Conditions PUF
2013	DD condition questions dropped	Slight reduction in reporting of conditions associated with missed work or school (e.g., "Cold," "Flu").	**MISSWORK, MISSSCHL, INBEDFLG** dropped from Conditions PUF
2008	PG section dropped	Little to no impact on reported conditions.	None
2007	PE section added	Increase in reporting of priority conditions (both treated and untreated).	**CONDBEG[M/D/Y]** variables replaced with **AGEDIAG** in Conditions PUF **REMISSN** added for cancer conditions (2007–2012) in Conditions PUF
2002	DN condition questions dropped	Slight reduction in reported conditions related to dental care (e.g., "Chipped tooth"), including reduction in treated conditions.	**DNNUM** dropped from Conditions PUF
1999	AP condition questions dropped	Slight reduction in reported conditions associated with alternative care (e.g., "Stress").	**APCARE3/APCARE53** dropped from Conditions PUF
1999	LC section dropped	Slight reduction in reported conditions associated with disabilities (e.g., "Old age").	None

Note. AP = Alternative/Preventive Care; CE = Condition Enumeration; DD = Disability Days; DN = Dental Care; ICD-9 = International Classification of Diseases, Ninth Revision; ICD-10 = International Classification of Diseases, Tenth Revision; LC = Long-term Care; PE = Priority Conditions Enumeration; PG = Pregnancy Detail; PUF = public-use file.

The following sections provide additional details for each of these changes. Each section includes a description of the change, commonly reported conditions (where applicable), variables affected, and the potential impact on conditions reported in the Medical Conditions PUFs. Tables include weighted counts of the total number of records in the Medical Conditions files, as well as counts limited to *treated* conditions and *treated prevalence*. Treated conditions are conditions that were discovered or led to a medical event or prescribed medicine purchase. Treated prevalence is defined as the number of people reporting a condition or group of conditions. Note that the number of treated conditions may be greater than treated prevalence due to the possibility of people with "duplicate" condition records in the MEPS Medical Conditions file (as discussed previously) and because similar conditions are grouped in the tables for expository purposes.

Please note that while this document covers the major changes to MEPS data collection and processing that affect the Medical Conditions PUFs, additional minor changes to survey questions over the years have affected the variables included in the MEPS Medical Conditions PUFs. A full list of variables for each Medical Conditions PUF over the years can be found by using the MEPS-HC Variable Explorer Tool at the following web page: https://datatools.ahrq.gov/meps-hc#varexpLabel.

2020: Conditions Pick-List Tool and Automated Probes

Description: Starting in 2020, a new conditions pick-list tool was added to the CAPI instrument to help interviewers record conditions associated with events. Previously, an interviewer would ask the respondent what the conditions were and then type those conditions into the survey instrument as verbatim text. Afterwards, the verbatim text would be matched to a historical list of condition text strings with the associated ICD-9 or ICD-10 codes. Any codes that did not map to a condition on this list would be sent to trained coders who would then assign the appropriate ICD-9 or ICD-10 code.

Starting in 2020, the Condition Pick-List tool allowed interviewers to search for conditions from a prespecified list and select the appropriate condition, rather than having to type it in manually. (If a condition does not appear in the list, the interviewer still has the option to manually enter the condition). In addition, the tool also flags commonly reported procedures or symptoms, indicating to the interviewer that they need to probe for the condition causing the procedure or symptom. For instance, if the interviewer types "SURGERY," the pick-list tool will indicate that the interviewer should probe for the underlying condition that necessitated the surgery. Prior to the implementation of the pick-list tool, interviewers were trained to probe for information in these scenarios. However, the pick-list tool automates this process by flagging cases that need additional probing in real time. This reduces human error because interviewers are not relying on their memory or judgment to determine which reported conditions require additional probing. The introduction of the pick-list tool significantly reduced the number of conditions sent to trained coders by 72.8 percent in 2020 compared with 2019.

Variables affected: No variables in the Medical Conditions PUFs were affected by this change.

Impact on number of records: This change likely has little to no impact on the number of condition records. However, analyses indicate that the pick-list tool and its automated probes slightly reduce the number of "uncodeable" conditions, as well as ambiguous text strings (e.g., "flu" could mean "stomach flu" or "respiratory influenza").

2018: Condition Enumeration (CE) Section Dropped

Description: From 1996 to 2016, respondents were asked a general question in the CE section of MEPS to identify any specific physical or mental health problems encountered by household members that bothered them during the interview reference period. These questions were asked prior to the sections on healthcare use or disability days, with the intent of priming the respondents' memories to better recall medical visits or disability days. The goal of asking these questions in the CE section was not to gain a comprehensive listing of all medical conditions. As a result, the condition responses cannot be used to estimate true medical condition prevalence. Starting with interviews conducted in 2018, this section was dropped. Consequently, in files for 2018 and beyond, all medical conditions records are those related to medical events or prescribed medicine purchases.

Conditions commonly reported only in the CE section: Common cold, depression, flu, allergies

Variables affected: No variables in the Medical Conditions PUFs were affected by this change.

Impact on number of records: This change resulted in a notable reduction in the number of total condition file records beginning in 2018 relative to earlier years (977.2 million weighted records in 2018 vs. 1.2 billion in 2017) but it had little impact on the number of treated conditions (see Table 1). In the earlier years of the survey (1996–2006), 40 to 50 percent of all condition records (weighted) were generated from questions asked in the CE section. After the introduction of the PE section in 2007, this number dropped to around 30 to 35 percent, because a portion of the conditions that had been generated from the CE section were now being generated from the PE section. Of those conditions that originated in the CE section, about 40 to 45 percent (11 to 15 percent of the total conditions records) were *not* associated with a medical event or prescribed medicine purchase, meaning that these events would not be captured in the ME sections of the survey. Thus, we see an expected decrease in the number of records in the Medical Conditions files because of the removal of this section from the survey.

This drop is also evident in the decrease in reported conditions for which a respondent may not see a doctor, such as a stomach bug, stress, or the common cold, as shown in Table 3. Note that while the total number of reported conditions decreased from 2016 to 2018, the number of *treated* conditions and treated prevalence was not affected by the removal of the CE section. This emphasizes the importance of limiting analyses to treated conditions, especially when viewing conditions before and after 2017.

Table 3. Weighted Number of Condition Records and People (in Millions) for Selected ICD10CDX Values, 2016–2019

ICD10CDX	Conditions	2016*	2017^	2018	2019
A08, R10	Stomach bug/stomachache				
	All reported conditions	12.2	11.1	5.8	5.7
	Treated conditions	6.2	6.0	5.8	5.7
	Treated prevalence	6.0	5.8	5.5	5.6
F32, F41, F43	Depression, anxiety, stress				
	All reported conditions	68.2	65.0	50.8	53.2
	Treated conditions	49.2	48.4	50.8	53.2
	Treated prevalence[1]	37.3	37.2	38.9	40.0

ICD10CDX	Conditions	2016*	2017^	2018	2019
J00, J11	Common cold, Flu				
	All reported conditions	54.5	50.5	20.7	17.5
	Treated conditions	16.6	19.2	20.7	17.5
	Treated prevalence	16.0	18.6	20.0	16.8

*Cells with green shading indicate years in which the Condition Enumeration (CE) section was fielded.
^ 2017 was a transition year for the Condition Enumeration (CE) section because CE information was collected for part of the 2017 data year (Rounds 3 and 4 of Panel 21 and Rounds 1 and 2 of Panel 22) and these conditions were included on the 2017 Medical Conditions public-use file.
[1] Treated prevalence is defined as the number of unique people reporting at least one of the conditions indicated by ICD10CDX. For instance, a person reporting a stomach bug (A08) and a stomachache (R10) is only counted once.

2016: New Condition Classification Coding (ICD-9 to ICD-10)

Description: Medical conditions reported by MEPS respondents are later mapped by professional coders to ICD-10 diagnosis codes, which are then collapsed into a smaller number of clinically meaningful categories called the CCSR codes; these codes are created and maintained by AHRQ's Healthcare Cost and Utilization Project (HCUP). As of version v2023.1 of the CCSR for ICD-10-CM, each ICD-10 code can map to up to six CCSR codes. This one-to-many mapping allows the CCSR codes to reflect the detailed nature of ICD-10 codes, which can describe multiple conditions, or a condition with a common symptom or manifestation (see example below). Most of the conditions reported in MEPS are coded to ICD-10 codes that map to three CCSR codes, which are released in the Medical Conditions PUFs in the variables CCSR1X, CCSR2X, and CCSR3X. In the 2016–2020 data files, 92–95 percent of reported conditions mapped to one CCSR code, 5–8 percent mapped to two CCSR codes, and fewer than 0.1 percent mapped to three or more CCSR codes. For the few cases that map to more than three CCSR codes, the extra CCSRs are zero-filled and collapsed. To preserve confidentiality of the PUFs, the ICD-10 codes are collapsed from the fully specified versions into three-digit top-codes and released in the PUFs in the variable ICD10CDX.

Note that the CCSR variables are listed in alphabetical order in the MEPS files, not in order of importance. For instance, the reported condition of "STREP THROAT" is coded to an ICD-10 value of J02.0, with the following ICD10CDX and CCSR pattern in the 2020 MEPS Conditions file (the value of "-1" is a reserve code for "Inapplicable"):

ICD10CDX	CCSR1X	CCSR2X	CCSR3X
J02: Acute Pharyngitis	INF003: Bacterial infections	RSP006: Other specified upper respiratory infections	-1

In this example, INF003 is listed under the CCSR1X variable, while RSP006 is listed under CCSR2X, even though RSP006 would be considered the "default" code for the ICD10CDX value of J02 (based on version v2023.1 of the HCUP CCSR crosswalk).

For the data years 1996–2015, a previous version of ICD diagnosis codes was used to classify reported conditions. These ICD-9 codes were then collapsed into Clinical Classifications Software (CCS) codes, a predecessor to the CCSR codes. In general, ICD-10 codes are more detailed than ICD-9 codes. For instance, while there are more than 14,000 ICD-9 diagnosis codes, there are nearly 70,000 ICD-10

diagnosis codes. In addition, ICD-10 codes can have up to seven characters, while ICD-9 codes have a maximum of five characters. Similar to the ICD-10 codes, these fully specified ICD-9 codes are collapsed into the three-digit top codes on the Medical Conditions PUFs (ICD9CODX).

Unlike the ICD-10 and CCSR mapping, each ICD-9 code maps to exactly one CCS code. For instance, "STREP THROAT" reported prior to 2016 was coded to:

ICD-9	CCCODEX
034.0: Strep sore throat	126: Other upper respiratory infections

Another difference between the current files and the 1996–2015 files is the treatment of procedures. In the current MEPS coding process, when a respondent reports a procedure instead of a condition, the procedure is coded to the underlying condition if it can be discerned from the text (e.g., "cataract surgery" is coded to ICD10CDX for "H26: Other Cataract"). In previous years (1996–2015), however, the procedures were included in the Medical Conditions PUF as a separate variable (ICD9PROX).

Note that ICD-9, ICD-10, CCS, and CCSR codes are subject to change over time. For instance, in 2004, the CCS codes related to mental disorders were revised, and in 2020, ICD-10 and CCSR codes related to COVID-19 were added. Additional information about ICD-9, ICD-10, CCS, and CCSR codes can be found at the following links:

- ICD-9 codes (CDC): https://cdc.gov/nchs/icd/icd9cm.htm
- ICD-10 codes (CDC): https://cdc.gov/nchs/icd/icd-10-cm.htm.
- CCS codes (HCUP): https://hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp
- CCSR codes (HCUP): https://hcup-us.ahrq.gov/toolssoftware/ccsr/ccs_refined.jsp

2016–2017 File Re-Release: The CCSR codes were not included in the initial release of the 2016 and 2017 Medical Conditions files because the ICD10-CCSR crosswalk was still being developed when the files were released. These variables were added and the files were subsequently re-released in the fall of 2022.

Prior to releasing the Medical Conditions data as public-use files, rare conditions are "masked" to preserve confidentiality. This masking process will either set the ICD-10 value to missing ("-15"), or re-code it into a more generic condition code. When condition coding transitioned from ICD-9 to ICD-10 codes, the residual masking processes had inadvertent effects on the 2016 and 2017 Medical Conditions PUFs, resulting in several conditions (ICD-10 codes) being over-masked in these files. As an example, masking processes that were carried over into 2016 and 2017 had the residual effect of re-coding E55 and G47 to the "broader" codes of E63 and G98, respectively. Table 4 presents the number of conditions (unweighted) from the initial and re-released Medical Conditions PUFs for some of the conditions that were initially over-masked.

These masking processes were reviewed and updated for subsequent data files, and the over-masked ICD10CDX values on the 2016 and 2017 files were edited and files were re-released in the fall of 2022.

Table 4. Number of Conditions (Unweighted) in the Initial and Re-Released 2016 and 2017 PUFs

ICD-10 Condition	Initial Release		Re-Release	
	2016	2017	2016	2017
E55* VITAMIN D DEFICIENCY	0	0	369	362
E63 OTHER NUTRITIONAL DEFICIENCIES	372	365	3	3
G47* SLEEP DISORDERS	0	0	1,723	1,662
G98 OTHER DISORDERS OF NERVOUS SYSTEM NEC	1,799	1,738	14	16
H26 OTHER CATARACT	0	0	635	665
I51 COMPLICATIONS AND ILL-DEFINED DESCRIPTIONS OF HEART DISEASE	0	0	361	338
M06 OTHER RHEUMATOID ARTHRITIS	0	641	727	641

ICD-10 = International Classification of Diseases, Tenth Revision; PUF = public-use file.
* Initial masking processes for the 2016 and 2017 data files re-coded the ICD-10 codes of E55 and G47 into the broader codes of E63 and G98, respectively.

Variables affected: The variables ICD9CODX, ICD9PROX, and CCCODEX were dropped starting with the 2016 file. Starting with the 2016 PUFs, ICD10CDX (the 3-digit ICD-10 top code), CCSR1X, CCSR2X, and CCSR3X variables were available. In addition, starting in 2016, CCS codes (CCCODEX) were no longer available on the event files. Instead, users must link event and conditions files using the event-condition linkage file (CLNK file). Table 5 shows the variables available for each data year.

Table 5. ICD and CCS/CCSR Variables by Year and Panel

Panel #	2014	2015	2016	2017	2018
Panel 19	ICD9CODX ICD9PROX CCCODEX	ICD9CODX ICD9PROX CCCODEX			
Panel 20		ICD9CODX ICD9PROX CCCODEX	ICD10CDX CCSR1X* CCSR2X* CCSR3X*		
Panel 21			ICD10CDX CCSR1X* CCSR2X* CCSR3X*	ICD10CDX CCSR1X* CCSR2X* CCSR3X*	
Panel 22				ICD10CDX CCSR1X* CCSR2X* CCSR3X*	ICD10CDX CCSR1X CCSR2X CCSR3X
Panel 23					ICD10CDX CCSR1X CCSR2X CCSR3X

CCS = Clinical Classification Software; CCSR = Clinical Classification Software Refined; ICD = International Classification of Diseases.
* Included in the re-released files

Impact on number of records: While the transition from ICD-9 to ICD-10 did not have a direct impact on the number of records in the Medical Conditions file, users should take care when using CCSR codes in analyses. The multiplicity of the CCSR codes provides an advantage relative to the CCS codes because they allow for more detail and accuracy when mapping the ICD-10 codes to the broader CCSR categories. As described in the "STREP THROAT" example above, the ICD-10 code maps to two CCSR categories (**INF003: Bacterial infections** and **RSP006: Other specified upper respiratory infections**), while the ICD-9 code maps to a single CCS category (126: Other upper respiratory infections). Due to these fundamental differences in the coding of ICD and CCS(R) codes, analysts are urged to use extreme caution when comparing conditions based on ICD-10 codes (2016 and later) with those based on ICD-9 codes (1996–2015).

2013: Disability Days (DD) Condition Questions Dropped

Description: Prior to 2013, household respondents were asked to report conditions that caused sample people to miss school or work or spend more than half a day in bed. Questions that asked about medical conditions associated with disability days were dropped in 2013 to reduce respondent burden, and because the data were not analytically useful when people had multiple conditions causing disability days.

Commonly reported conditions: Cold, influenza, stomach virus/flu, headache, fever, sore throat

Variables affected: MISSWORK, MISSSCHL, and INBEDFLG, which had been derived from the DD section, are not included in the Medical Conditions files for 2013 onward. Table 6 displays the variables available in the PUFs by data year and round.

Table 6. MISSWORK, MISSSCHL, and INBEDFLG Variables by Panel, Year, and Rounds

Panel #	2012	2013	2014
Panel 16	Rounds 3–5 MISSWORK MISSSCHL INBEDFLG		
Panel 17	Rounds 1–3 MISSWORK MISSSCHL INBEDFLG	Rounds 3–5 -- -- --	
Panel 18		Rounds 1–3 -- -- --	Rounds 3–5 -- -- --

Impact on number of records: Around 5–8 percent of condition records (weighted) were generated from the DD section prior to its removal (Table 1). The removal of the condition questions in the DD section reduced the total number of conditions that were reported but it had little impact on treated conditions or treated prevalence (Table 7). The types of conditions that decreased are those that one can often treat at home without a medical provider (e.g., minor illnesses like "common cold" or "stomach flu").

Table 7. Weighted Number of Condition Records and People (in Millions) for Selected ICD9CODX Values, 2011–2015

ICD9CODX	Condition	2011*	2012*	2013^	2014	2015
460	Acute nasopharyngitis (e.g., common cold)					
	All reported conditions	44.7	46.0	44.4	31.5	34.8
	Treated conditions	9.7	9.2	9.5	9.4	9.7
	Treated prevalence	9.6	9.2	9.4	9.3	9.6
008, 487, 079	Other infections, influenza, viral infection (e.g., stomach virus/flu, flu, virus)					
	All reported conditions	47.4	50.0	52.2	13.7	12.5
	Treated conditions	12.4	12.1	13.8	6.2	6.1
	Treated prevalence	11.9	11.8	13.2	6.0	6.0

***** Cells with green shading indicate years in which the Disability Days (DD) condition questions were fielded.
^ 2013 was a transition year, where the DD condition questions were asked in the first part of the year only.

2008: Pregnancy Detail (PG) Section Dropped

Description: From 1996 to 2007, this section collected additional information for women identified as having been pregnant during the reference period or a previous round. Additional information included pregnancy outcome, delivery details, and complications experienced. Starting in 2008, this section was no longer fielded because it collected information about deliveries before questions about events, and some respondents did not report delivery events later in the survey, perhaps because respondents believed they had already reported delivery events. Eliminating the pregnancy detail section appears to have improved the quality of event and expenditure data.

Commonly reported conditions: Pregnancy

Variables affected: No variables on the Medical Conditions PUFs were affected by this change.

Impact on number of records: Only a small number of conditions (<1%) were reported in this section. In addition, pregnancy-related conditions are highly likely to be reported in the CE or ME sections. Thus, eliminating this section had minimal impact on the total number of records in the Medical Conditions PUF. However, as shown in Table 8, there was a slight increase in reports of "Normal" delivery (ICD9CODX=650), indicating that the removal of the PG section had the desired effect of encouraging respondents to report delivery events in the ME sections of the survey.

Table 8. Weighted Number of Conditions Records and People (in Millions) for Selected ICD9CODX Values, 2005–2010

ICD9CODX	Condition	2005*	2006*	2007*	2008	2009	2010
650	Normal delivery						
	All reported conditions	0.51	0.37	0.43	0.78	0.83	1.08
	Treated conditions	0.38	0.27	0.31	0.63	0.60	0.78
	Treated prevalence	0.38	0.27	0.31	0.63	0.60	0.76
V22	Normal pregnancy						
	All reported conditions	8.69	8.88	7.64	6.62	6.07	6.52
	Treated conditions	7.40	7.13	6.52	5.81	5.27	5.78
	Treated prevalence	6.73	6.47	5.95	5.62	5.19	5.65

***** Cells with green shading indicate years in which the Pregnancy Detail (PG) section was fielded.

2007: Priority Conditions Enumeration (PE) Section Dropped

Description: In Panel 12, which began in 2007, a redesigned MEPS-HC interview was fielded, and the Priority Conditions Enumeration (PE) section replaced the majority of the questions in the Priority Condition (PC) section. In both the PC and PE sections, the household respondents were asked a series of "yes/no" questions on whether each person had been diagnosed as having several specific conditions that are generally chronic in nature. These conditions were classified by AHRQ as "priority" due to their relatively high prevalence, their priority for the Agency, or because generally accepted standards for appropriate clinical care have been developed. Responses to these priority condition questions are recorded in the Full-Year Consolidated PUF. These conditions, along with their corresponding variable names (in the Full-Year Consolidated file) and three-digit ICD-9/ICD-10 codes (on the Medical Conditions files) are listed in Table 9.

Table 9. Priority Conditions, Panel 12 and After, With Corresponding Variable Name (for Full-Year Consolidated File) and ICD9CODX and ICD10CDX Values (for Medical Conditions File)

Priority Condition	Variable Name in FYC File	ICD9CODX (2008–2015)	ICD10CDX (2016 and Later)
Angina/Angina Pectoris	ANGIDX	413	I20
Arthritis – Osteoarthritis	ARTHDX	715	M19
Arthritis – Rheumatoid Arthritis	ARTHDX	714	M06
Arthritis (not specified)	ARTHDX	716	M19
Asthma	ASTHDX	493	J45
Attention Deficit Hyperactivity Disorder (ADHD) / Attention Deficit Disorder (ADD)	ADHDADDX	314	F90
Cancer/Malignancy	CANCERDX	140-239[1]	C00-C97 D00-D49[1]
Chronic Bronchitis	CHBRON31	491	J42
Coronary Heart Disease	CHDDX	414	I25
Diabetes/Sugar Diabetes	DIABDX (2007-2017) DIABDX_M18 (2018-current)	250	E11
Emphysema	EMPHDX	492	J43
Heart Attack/Myocardial Infarction (MI)	MIDX	410	I21
High Cholesterol	CHOLDX	272	E78
Hypertension/High Blood Pressure	HIBPDX	401	I10
Joint Pain	JTPAIN31 (2007-2017) JTPAIN31_M18 (2018-current)	719	M25
Other Heart Disease (not coronary heart disease, angina, or heart attack)	OHRTDX	[multiple][1]	[multiple][1]
Stroke / Transient Ischemic Attack (TIA) / Mini stroke	STRKDX	436	G45

ICD-9 = International Classification of Diseases, Ninth Revision; ICD-10 = International Classification of Diseases,Tenth Revision.

[1] Respondents answering "Yes" to having cancer or other heart disease were then asked follow-up questions on the specific type of cancer or heart disease. These responses were then coded into various ICD-9 or ICD-10 codes.

Starting in Panel 12, when a respondent answers "yes" to a priority condition question in the PE section, that condition is automatically added to the dynamic conditions roster for the person. A record for that condition is then added to the Medical Conditions file if it is also reported in the ME section (or in the CE section pre-2018; or in the DD section pre-2013).

In addition to the major change in how priority conditions are recorded, the list of priority conditions generally expanded in Panel 12, and the variable AGEDIAG was added in 2007 to identify age of diagnosis for all priority conditions (excluding joint pain).

For cancer conditions collected in the PE section, follow-up questions were asked to identify the type of cancer, as well as when the cancer was first reported to determine whether the cancer was in remission/under control (REMISSN). Starting in the 2013 PUF, two changes were made to the PUF to maintain respondent confidentiality: REMISSN was dropped and the AGEDIAG variable was reset to -1 for all cancer records.

Variables affected: AGEDIAG was added in 2007, while CONDBEGD, CONDBEGM, and CONDBEGY were dropped from the 2008 file and all files after 2008.The REMISSN variable specific to cancer priority conditions is available in the 2007–2012 Medical Conditions PUFs. Table 10 shows the affected variables by panel and year. Note that in 2007, AGEDIAG and REMISSN = -1 (Inapplicable) for Panel 11, while the CONDBEG[M/D/Y] variables = -1 (Inapplicable) for Panel 12.

Table 10. CONDBEG[M/D/Y] and AGEDIAG Variables by Year and Panel

Panel #	2006	2007	2008
Panel 11	CONDBEG[M/D/Y]	CONDBEG[M/D/Y] AGEDIAG = -1 REMISSN = -1	
Panel 12		CONDBEG[M/D/Y] = -1 AGEDIAG REMISSN	AGEDIAG REMISSN
Panel 13			AGEDIAG REMISSN

Impact on number of records: Because the PE section asks questions about specific conditions, respondents notably reported more priority conditions (for overall conditions as well as treated conditions) in the years following the addition of the PE section. Table 11 provides the weighted number of total reported conditions, treated conditions, and treated prevalence for the ICD-9 codes corresponding to the priority conditions (with the exception of "Other heart conditions," due to the wide variety of codes in this category). Because of the increased reporting of priority conditions after the redesign, we recommend limiting analyses of priority conditions to 2008 and later.

Table 11. Weighted Number of Condition Records and People (in Millions) for ICD-9 Codes Associated With Priority Conditions, 2005–2010

ICD9CODX	Priority Condition	2005	2006	2007^	2008*	2009*	2010*
314	ADHD/ADD						
	All reported conditions	4.4	4.3	5.0	6.2	6.8	7.2
	Treated conditions	4.0	4.1	4.4	5.3	5.8	5.9
	Treated prevalence	3.9	4.0	4.3	5.1	5.5	5.7
413	Angina/Angina Pectoris						
	All reported conditions	1.1	0.9	2.6	4.3	3.8	3.5
	Treated conditions	0.8	0.8	2.2	3.2	2.8	2.7
	Treated prevalence	0.8	0.8	2.1	3.1	2.7	2.6
714-716	Arthritis						
	All reported conditions	22.4	21.6	29.2	36.5	36.0	37.1
	Treated conditions	16.3	15.7	19.8	24.6	24.3	26.1
	Treated prevalence	14.9	14.4	18.4	22.7	22.6	24.0
493	Asthma						
	All reported conditions	15.7	16.8	17.6	19.7	19.9	21.0
	Treated conditions	12.4	13.6	13.3	15.1	15.3	16.4
	Treated prevalence	12.1	13.2	13.1	14.8	15.0	16.1
140-239	Cancer						
	All reported conditions	18.9	19.8	22.8	25.6	24.8	24.9
	Treated conditions	16.2	17.1	19.3	22.1	21.1	21.1
	Treated prevalence	14.0	14.7	16.2	18.4	17.7	17.8
491	Chronic Bronchitis						
	All reported conditions	0.5	0.6	2.8	5.9	5.0	5.2
	Treated conditions	0.4	0.3	2.0	4.3	3.5	3.5
	Treated prevalence	0.4	0.3	2.0	4.3	3.5	3.5
414	Coronary Heart Disease						
	All reported conditions	1.2	1.1	5.6	12.8	12.3	12.2
	Treated conditions	1.0	1.0	5.2	11.6	11.1	11.3
	Treated prevalence	1.0	1.0	4.9	10.9	10.5	10.6
250	Diabetes/Sugar Diabetes						
	All reported conditions	18.3	19.4	21.3	22.7	22.3	24.0
	Treated conditions	17.3	18.6	19.9	21.6	21.0	22.6
	Treated prevalence	16.5	17.6	18.9	20.7	20.0	21.5
492	Emphysema						
	All reported conditions	1.3	1.5	2.2	3.6	3.4	3.5
	Treated conditions	1.1	1.4	1.9	2.8	2.7	2.8
	Treated prevalence	1.1	1.3	1.9	2.8	2.7	2.7
410	Heart Attack / Myocardial Infarction (MI)						
	All reported conditions	1.8	1.7	3.5	6.1	6.3	6.3
	Treated conditions	1.6	1.5	3.0	5.3	5.5	5.6
	Treated prevalence	1.5	1.5	2.9	5.0	5.2	5.2

ICD9CODX	Priority Condition	2005	2006	2007^	2008*	2009*	2010*
272	High Cholesterol						
	All reported conditions	31.6	33.2	42.6	53.1	52.6	53.2
	Treated conditions	29.3	30.7	37.9	46.3	46.8	47.9
	Treated prevalence	28.6	30.2	36.9	44.5	45.2	46.1
401	Hypertension / High Blood Pressure						
	All reported conditions	48.6	49.2	55.5	61.6	62.4	64.3
	Treated conditions	46.0	46.6	51.6	57.2	58.0	60.3
	Treated prevalence	45.2	45.8	50.5	55.2	56.5	58.7
719	Joint Pain						
	All reported conditions	12.6	12.8	19.1	29.6	31.8	32.4
	Treated conditions	8.7	9.0	11.8	16.5	17.1	18.8
	Treated prevalence	8.0	8.2	10.9	15.3	15.9	17.0
436	Stroke/Transient Ischemic Attack (TIA)/ Mini stroke						
	All reported conditions	1.5	1.5	3.1	4.3	4.3	4.5
	Treated conditions	1.2	1.2	2.5	3.4	3.5	3.4
	Treated prevalence	1.2	1.1	2.4	3.3	3.4	3.4

ADD = attention deficit disorder; ADHD = attention-deficit/hyperactivity disorder; ICD-9 = International Classification of Diseases, Ninth Revision.
^ 2007 is a transition year, where the old version of the survey was fielded to Panel 11, and the new version with the added Priority Conditions Enumeration (PE) section was fielded to Panel 12.
***** Cells with green shading indicate years in which the revised PE section was fielded.

2002: Dental Care (DN) Condition Questions Dropped

Description: From 1996 to 2001, the Dental Care section of the MEPS-HC survey asked whether the dental care was related to an accident or injury, and if so, what kind of injury. The description of the injury was added to the conditions roster. Starting with interviews for Round 3 of Panel 6 and Round 1 of Panel 7 (conducted in the spring of 2002), these questions were dropped from the Dental Care survey section because it was burdensome to ask the questions for every visit, and the questions yielded so few conditions.

Commonly reported conditions: Broken/chipped tooth, broken/cracked filling

Variables affected: Starting in 2001, the variable DNNUM was no longer included in the Medical Conditions PUF. This variable is excluded from the 2001 file because the condition question was dropped from the Dental Care section for Round 3 of Panel 6, which includes data on dental visits from both 2001 and 2002.

Impact on number of records: Only a small number of conditions (<1%) were reported in this section. Thus, eliminating the questions from this section likely had minimal impact on the total number of records in the Medical Conditions PUF. However, there may be a decrease for a small number of specific conditions, particularly those that were commonly reported in the DN section but would be less likely to be captured in other survey sections (e.g., CE or other ME sections). For instance, “broken tooth” or “chipped tooth” (part of broader category ICD9CODX = 873) display a slight drop-off in 2002, after the DN condition questions were dropped from the survey (Table 12). Note that this decrease is also evident when looking only at *treated* conditions, because this survey change removed the link between conditions related to dental events.

Table 12. Weighted Number of Conditions Records and People (in Millions) for ICD9CODX = 873, 1999–2003

ICD9CODX	Condition	1999*	2000*	2001*	2002	2003
873	Other open wound of head (e.g., broken/chipped tooth)					
	All reported conditions	4.1	4.3	4.4	3.1	2.7
	Treated conditions	3.3	3.4	2.0	1.6	1.6
	Treated prevalence	3.2	3.3	2.0	1.6	1.5

***** Cells with green shading indicate years in which the Dental Care (DN) condition questions were fielded.

1999: Alternative/Preventive Care (AP) Condition Questions Dropped

Description: This section was a supplemental section that gathered information on any preventive care. Originally, this section included questions related to preventive care and alternative and complementary care, including the question "For what health problems was the alternative care practitioner consulted?" The section was not fielded in 1999.

Commonly reported conditions: Stress, weight loss, back pain

Variables affected: APCARE3 and APCARE53 are only available on the 1996 and 1998 Medical Conditions PUFs, respectively. These variables denote whether the respondent received alternative care for a condition. Table 13 shows the variables available in each data year.

Table 13. APCARE Variable by Year and Panel

Panel #	1996	1997	1998	1999
Panel 1	APCARE3 = 0,1,2	--		
Panel 2		--	APCARE53 = -1, 0, 1, 2	
Panel 3			APCARE53 = -1, 0, 1, 2	--

Impact on number of records: Only a small number of conditions (<1%) were reported in this section. Thus, eliminating these questions had minimal impact on the total number of records in the Medical Conditions PUF. However, there may be a decrease for a small number of specific conditions, particularly those that were commonly reported in the AP section but would be less likely to be captured in other survey sections (e.g., CE or ME sections). For instance, records with ICD9CODX = 308 (e.g., "Stress") display a slight drop off among all reported conditions in 1999 and 2000, after the conditions question from the AP section were dropped from the survey (Table 14). Note, however, that the same decline is not evident when looking at treated conditions or treated prevalence.

Table 14. Weighted Number of Condition Records and People (in Millions) for ICD9CODX = 308, 1996–2000

ICD9CODX	Conditions	1996	1997*	1998*	1999	2000
308	Acute reaction to stress					
	All reported conditions	5.5	5.4	5.1	4.2	3.6
	Treated conditions	1.8	2.4	2.4	2.1	1.6
	Treated prevalence	1.7	2.3	2.4	2.1	1.6

* Cells with green shading indicate years in which the Alternative/Preventive Care (AP) section was fielded.

1999: Long-Term Care (LC) Section Dropped

Description: This section was a supplemental section fielded in 1997 and 1998 that collected detailed information on individuals who have activity or functional limitations, hearing or vision impairments, and special equipment requirements. Information was also collected on conditions associated with the limitations.

Commonly reported conditions: Old age, learning disability, stroke, ADD

Variables affected: No variables in the Medical Conditions PUFs were affected by this change.

Impact on number of records: Only a small number of conditions (<1%) were reported in this section. Thus, eliminating this section likely had minimal impact on the total number of records in the Medical Conditions PUF. However, this change may have resulted in decreased reporting of a small number of specific conditions, particularly those that were commonly reported in the LC section but would be less likely to be captured in other survey sections (e.g., CE or ME sections). For instance, records with ICD9CODX = 797 (e.g., "Old Age") or 315 (e.g. "Learning disability") both slightly increased in 1997 and 1998, when the long-term care section was fielded, and decreased in 1999 after it was dropped (Table 15). Note, however, that the same trends are not as noticeable when looking only at treated conditions or treated prevalence. In fact, the weighted number of treated condition records for ICD9CODX = 315 increased slightly in 1999 (from 0.13 million to over 0.22 million).

Table 15. Weighted Number of Condition Records and People (in Millions) for Selected ICD9CODX values, 1996–2000

ICD9CODX	Conditions	1996	1997*	1998*	1999	2000
315	Specific developmental delays (e.g., learning disability)					
	All reported conditions	0.24	0.51	0.52	0.28	0.28
	Treated conditions	0.19	0.13	0.13	0.22	0.20
	Treated prevalence	0.19	0.12	0.13	0.22	0.20
797	Senility without psychosis (e.g., old age)					
	All reported conditions	0.91	1.20	1.04	0.69	0.67
	Treated conditions	0.85	0.85	0.78	0.66	0.61
	Treated prevalence	0.85	0.83	0.74	0.66	0.57

* Cells with green shading indicate years in which the Long-term Care (LC) section was fielded.

Other Changes

Description: In addition to the changes to the MEPS Medical Conditions PUFs described above, various other changes to the survey questionnaire over the years have resulted in additional variables being removed or edited in the Medical Conditions files. Most of these changes were a result of survey changes implemented to improve reporting and reduce respondent burden by eliminating survey questions asking about condition-specific details.

Additional improvements to the entire survey may also have affected the reporting of conditions. For instance, major changes to the survey were implemented in 2013 and 2018 to improve event reporting (Zuvekas, Biener, and Hicks, 2020). This improvement may have subsequently affected the number of events linked to medical conditions, and thus indirectly improved reporting of treated conditions.

Variables affected: The following variables were affected in the Medical Conditions PUF due to survey design changes.

Table 16. Variables in the Medical Conditions File Affected by Other Survey Changes

Year	Change
2001	HSNUM renamed IPNUM
2005	Variable dropped: WHOTYP#: Who reported condition [person with condition vs. another household member]?
2007	Variables dropped: GUN: Was a gun involved [in the accident/injury]? LSTSAW1: When was the last time doctor was seen?
2008	Condition detail variables dropped: FOLOCA1#: Received follow-up care for condition FURTCA1#: Further treatment recommended OVRALL1#: How condition affected overall health SEEDREF#: Saw doctor in reference period STILTR#: Is person still treated for condition?
2008	Accident/injury follow-up questions dropped: ACDNTLOC: Where did accident happen? ACDNTOTH: Was something else involved? INJURFLG: Location of round-specific injury info INOUTHH: Was accident inside/outside the house? PRIORFLG: Location of round-specific priority info RECOVER: Fully recovered from condition? DROWN: Was drowning/near-drowning involved? FALL: Was it a fall? FIREBURN: Was fire/burning involved? POISON: Was poison/poisonous substance involved? SPORTS: Was it a sports injury? VEHICLE: Was a motor vehicle involved? WEAPON: Was some other weapon involved?
2010	PRIOLIST variable dropped

Year	Change
2013	Variables dropped: ACCDENT[M/D/Y]: Day/month/year accident or injury occurred ACCDNJAN: Whether accident or injury occurred before or after Jan 1
2018	Information on office-based or outpatient "phone calls" no longer collected: OBNUM and OPNUM no longer include "phone calls"
2020	Telehealth questions added to survey: OBNUM and OPNUM include telehealth visits

Note: # = round number. For instance, "WHOTYP#" includes WHOTYP1–WHOTYP5 variables.

Conclusion

MEPS data are a useful resource for examining condition-specific treatment and expenditures in the U.S. civilian noninstitutionalized population. This resource becomes even more powerful when used to examine changes over time to understand a more complete picture of underlying trends in treated or diagnosed conditions and healthcare. However, analysts should be aware that the modifications to the MEPS instrument design and condition coding over the years as described in this document have the potential to affect trend analyses.

Changes in the instrument over the years have generally been designed to make the data collection effort more efficient and easier to administer, thus reducing respondent burden. For instance, the primary goal of recent major design improvements in 2018 was to reduce underreporting of healthcare events and improve data quality. Process changes, such as data editing, and changes in coding and imputation may also affect trend analyses.

This document details changes to MEPS data collection and processing that have affected how medical conditions are collected, processed, and reported in the Medical Conditions PUFs. Users are encouraged to use caution when analyzing condition data over time, because apparent differences in condition estimates could merely be due to changes in survey design and administration or data processing. Thus, in addition to considering external factors such as ICD coding changes, healthcare policies, and global events (e.g., the COVID-19 pandemic), any trend analyses using MEPS data should also carefully consider the potential impact of changes to the survey administration or data processing.

References

Cohen, Joel. 1997. Design and Methods of the Medical Expenditure Panel Survey Household Component. MEPS Methodology Report No. 1. AHCPR Pub. No. 97-0026. Rockville, MD: Agency for Healthcare Policy and Research. http://www.meps.ahrq.gov/mepsweb/data_files/publications/mr1/mr1.pdf

Ezzati-Rice, Trena M., Frederick Rohde, and Janet Greenblatt. 2007. Sample Design of the Medical Expenditure Panel Survey Household Component, 1998–2007. Methodology Report No. 22, Agency for Healthcare Research and Quality. http://www.meps.ahrq.gov/mepsweb/data_files/publications/mr22/mr22.pdf

Machlin, Steve R., Joel Cohen, Anne Elixhauser, Karen Beauregard, and Claudia Steiner. 2009. "Sensitivity of Household Reported Medical Conditions in the Medical Expenditure Panel Survey." Medical Care, 47(6): 618–625.

Stagnitti, Marie N., Steve R. Machlin, Marc W. Zodet, and Erica Saleska. Design, Methods, and Field Results of the Medical Expenditure Panel Survey Medical Provider Component (MEPS MPC) including the Medical Organizations Survey (MOS)— 2016 Data Year. Methodology Report #32. Agency for Healthcare Research and Quality. https://meps.ahrq.gov/data_files/publications/mr32/mr32.pdf

Zuvekas, Samuel H., Adam I. Biener, and Wendy D. Hicks. 2020. "The Effects of Survey Enhancements on the Quality of Reporting in the Medical Expenditure Panel Survey, 2008–2015." Journal of Survey Statistics and Methodology, 8(3): 589–616.

Appendices

Appendix A. Priority Condition Variables on the Full-Year Consolidated PUFs

Table A1. Priority Condition Variables Ever Diagnosed

Condition	Ever Diagnosed (1=Yes, 2=No)	Age of Diagnosis	Variables Derived From Follow-up Questions
High Blood Pressure (18+)	HIBPDX	HIBPAGED	BPMLDX - high blood pressure diagnosis on 2 or more visits
Heart Disease (18+)			
- Coronary Heart Disease	CHDDX	CHDAGED	
- Angina / Angina Pectoris	ANGIDX	ANGIAGED	
- Heart Attack / MI	MIDX	MIAGED	
- Other heart condition	OHRTDX	OHRTAGED	OHRTTYPE - Type of other heart condition (added in 2018) 1 = Heart murmur 2 = Heart arrhythmia/irregular heartbeat 3 = Blocked or clogged artery 4 = Congestive heart failure 5 = Atrial fibrillation 6 = Mitral valve prolapse 7 = Enlarged heart 8 = Heart valve problems 9 = Tachycardia/rapid heart rate 10 = Bradycardia/slow heart rate 91 = Other
Stroke/TIA (18+)	STRKDX	STRKAGED	
Emphysema (18+)	EMPHDX	EMPHAGED	
High cholesterol (18+)	CHOLDX	CHOLAGED	
Diabetes (all ages)	DIABDX_M18	DIABDAGE	
Arthritis (18+)	ARTHDX	ARTHAGED	ARTHTYPE – Type of arthritis 1 = Rheumatoid arthritis 2 = Osteoarthritis 3 = Not specified

Condition	Ever Diagnosed (1=Yes, 2=No)	Age of Diagnosis	Variables Derived From Follow-up Questions
Cancer (18+)	CANCERDX	(Dropped from PUFs in 2013)[1]	**Type of cancer[2]** CABLADDR – Bladder CABLOOD – Blood CABREAST – Breast CACERVIX – Cervical CACOLON – Colon CALUNG – Lung CALYMPH – Lymphoma (non-Hodgkin's) CAMELANO – Skin melanoma CAMUSCLE – Soft tissue, muscle or fat CAOTHER – Other type CAPROSTA – Prostate CASKINNM – Skin (non-melanoma) CASKINDK – Skin (unknown type) CAUTERUS – Uterine
Asthma (all ages)	ASTHDX	ASTHAGED	**Asthma follow-up questions asked in Rounds 3/1** ASSTIL31 - Does person still have asthma ASATAK31 - Asthma attack last 12 months ASTHEP31 - When was last episode of asthma ASACUT31 - Used acute pressure inhaler in last 3 months ⇨ ASMRCN31 - Used >3 canisters in last 3 months ASPREV31 - Ever used preventive daily asthma meds ⇨ ASDALY31 - Now take preventive daily asthma meds ASPKFL31 - Have peak flow meter at home ⇨ ASEVFL31 - Ever used peak flow meter ⇨ ASWNFL31 - When last used peak flow meter
ADHD (5-17)	ADHDADDX	ADHDAGED	

Note: ADHD=attention-deficit/hyperactivity disorder; MI = myocardial infarction; PUFs = public-use files; TIA = transient ischemic attack. People under age 18 will have a value of "-1 Inapplicable" for conditions asked only of adults.

[1] The "Age of Diagnosis" variables for each cancer type are available in the restricted data files at the AHRQ Data Center for 2013 and later.

[2] Some "Type of cancer" variables may not be available in the PUFs every year due to low prevalence.

Table A2. Priority Condition Variables Diagnosed in Last 12 months

Condition	Diagnosed in last 12 months (1=Yes, 2=No)	Notes
Chronic Bronchitis (18+)	CHBRON31	
Joint Pain (18+)	JTPAIN53_M18 (2020) JTPAIN31_M18 (2018 and later) JTPAIN31 (1996-2017) JTPAIN53 (1996-2016)	Starting in 2018, joint pain questions (JTPAIN31_M18) are skipped if person has already reported an arthritis condition.

Appendix B. Variables in the Medical Conditions PUFs

Survey design changes resulted in major variable changes to the Medial Conditions file starting in 2013.

Table B1. Variables in Medical Conditions Files for Data Years 2013 and later

Name	Data years (if partial)	Description
ACCDNWRK		DID ACCIDENT OCCUR AT WORK
AGEDIAG		AGE WHEN DIAGNOSED
CCCODEX	2015 and earlier	CLINICAL CLASSIFICAITON CODE - EDITED
CCSR1X	2016 and later	CLINICAL CLASSIFICATION REFINED CODE 1- EDITED
CCSR2X	2016 and later	CLINICAL CLASSIFICATION REFINED CODE 2- EDITED
CCSR3X	2016 and later	CLINICAL CLASSIFICATION REFINED CODE 3- EDITED
CONDIDX		CONDITION ID
CONDN		CONDITION NUMBER
CONDRN		CONDITION ROUND NUMBER
CRND1		HAS CONDITION INFORMATION IN ROUND 1
CRND2		HAS CONDITION INFORMATION IN ROUND 2
CRND3		HAS CONDITION INFORMATION IN ROUND 3
CRND4		HAS CONDITION INFORMATION IN ROUND 4
CRND5		HAS CONDITION INFORMATION IN ROUND 5
DUID		PANEL # + ENCRYPTED DU IDENTIFIER
DUPERSID		PERSON ID (DUID + PID)

Name	Data years (if partial)	Description
ERNUM		# ER EVENTS ASSOC. W/ CONDITION
HHNUM		# HOME HEALTH EVENTS ASSOC. W/ CONDITION
ICD10CDX	2016 and later	ICD-10-CM CODE FOR CONDITION - EDITED
ICD9CODX	2015 and earlier	ICD-9-CM CODE FOR CONDITION – EDITED
ICD9PROX	2015 and earlier	ICD-9-CM CODE FOR PROCEDURE - EDITED
INJURY		WAS CONDITION DUE TO ACCIDENT/INJURY
IPNUM		# INPATIENT EVENTS ASSOC. W/ CONDITION
OBNUM		# OFFICE-BASED EVENTS ASSOC W/ CONDITION
OPNUM		# OUTPATIENT EVENTS ASSOC. W/ CONDITION
PANEL		PANEL NUMBER
PERWT[yy]F^		EXPENDITURE FILE PERSON WEIGHT, 20[yy]^
PID		PERSON NUMBER
RXNUM		# PRESCRIBED MEDICINES ASSOC. W/ COND.
VARPSU		VARIANCE ESTIMATION PSU, [year]
VARSTR		VARIANCE ESTIMATION STRATUM, [year]

^ [yy] = 2-digit year (e.g., “PERWT[yy]F” = “PERWT20F” for data year 2020).