AGENCY FOR HEALTHCARE RESEARCH AND QUALITY

METHODOLOGY REPORT #36

February 24, 2023

Analyzing Medical Conditions in MEPS: User Guide (Part 1 of 2)

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

Summary

This report is the first 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 user guide contains specific guidance for analyzing MEPS condition data. Part 2 of this series is a detailed reference guide that documents changes to the survey that have affected how data on medical conditions are collected, processed, and disseminated to the public.

Analyzing Medical Conditions in MEPS: User Guide (Part 1 of 2)

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

Background

The Medical Expenditure Panel Survey Household Component (MEPS-HC) is a questionnaire that collects nationally representative data on healthcare use and expenditures from individual household members in the U.S. civilian noninstitutionalized population. The survey also collects information on medical conditions associated with healthcare events, such as hospital stays, physician visits, and prescription drug purchases. Each year, public-use files (PUFs) are released on the MEPS website (https://meps.ahrq.gov/) after data collection and processing are completed. Several of these PUFs contain data on medical conditions that can be used to analyze healthcare utilization and spending on medical conditions in the United States. For example:

- The **Full-Year Consolidated PUF** is a person-level file that contains data on demographics, healthcare spending, health status, and priority conditions. Most questions about priority conditions ask respondents whether adult household members were ever diagnosed with a particular condition. Follow-up questions such as age of diagnosis are also asked for certain conditions.

- The **Medical Conditions PUF** is a condition-level file, which contains a record for each reported 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 Classifications Software Refined (CCSR).[1] Condition records can be linked to the Full-Year Consolidated file and to the MEPS event files.

Guidance #1: Analyze treated conditions.

The MEPS focuses on healthcare service use and expenditures, so the best use of the Medical Conditions file is to analyze *treated* conditions, which are conditions associated with a healthcare event or prescribed medicine purchase. The MEPS Medical Conditions data do not provide estimates of true population prevalence.

For the most detailed analyses, and to identify treated conditions before 2018,[2] analysts should link the MEPS Medical Conditions file to the MEPS event files, using the Condition-Event Link (CLNK) file as a crosswalk; this will allow analysts to identify conditions that are associated with healthcare events or prescribed medicine purchases. These files have a many-to-many relationship because one condition can link to multiple events, and one event can link to multiple conditions. More information on linking the medical conditions file to the event files is included in the documentation for each CLNK file.[3] Example codes in R, SAS, and Stata can be found on the MEPS GitHub page at https://github.com/HHS-AHRQ/MEPS. There are six types of MEPS public-use event files that can link to medical conditions (see Table 1).

---

[1] The ICD-10-CM is available at https://cdc.gov/nchs/icd/icd-10-cm.htm. The CCSR is available at https://hcup-us.ahrq.gov/toolssoftware/ccsr/ccs_refined.jsp. For data years 1996–2015, previous versions of the ICD and CCS codes were released.

[2] For 2018 and later, all records in each MEPS Medical Conditions file are linked to a healthcare event or prescribed medicine purchase.

[3] Documentation for 2020 file: meps.ahrq.gov/data_stats/download_data/pufs/h220i/h220idoc.shtml.

Table 1. Summary of MEPS event files

Event Type	Each Record Is…	Notes
Hospital inpatient stays	A hospital inpatient stay	Each stay can consist of multiple nights in the hospital.
Emergency room visits	A visit to an emergency department	If a person was admitted to the hospital through the ER, the ER facility fees will be $0.
Outpatient visits	A visit to an outpatient department	For 1996–2017 data, informal telephone calls were included, with expenditures of "-1."
Office-based medical provider visits	A visit to an office-based medical provider	For 1996–2017 data, informal telephone calls were included, with expenditures of "-1."
Home health care	One month of care	Includes both paid and unpaid (informal) home health care. Expenditures for informal care have a value of "-1."
Prescribed medicines	A unique purchase or fill	An EVNTIDX value on the CLNK can represent multiple fills of a drug in a given round.

Caveats and notes:

- Conditions may link to medical events for which expenditure data were inapplicable, such as informal home health care or telephone calls for office-based medical provider or outpatient visits.[4] Expenditure variables in the event files are set to "-1" for these events. It is up to the user to decide whether to include these types of events when estimating "treated prevalence."

- For 1996–2017, MEPS respondents were asked about any medical conditions that bothered household members (in the "Condition Enumeration" section) before they were asked about healthcare use and disability days. The intent of this section was to prime respondents' memories to better recall medical visits, not to gain a comprehensive list of all medical conditions. Thus, this section cannot be used to capture true prevalence.

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

- The number of dental visits associated with an accident or injury (DNNUM) is available in data files from 1996 to 2000. However, the corresponding survey question was removed from the survey in 2001, and this variable was dropped from subsequent files.

Guidance #2: Use caution when comparing ICD-10 (2016 and later) and ICD-9 (1996–2015) conditions.

Medical conditions reported by MEPS respondents are coded by professional coders using ICD-10-CM codes, which are then mapped to CCSR codes. In the Medical Conditions PUFs for 2016 and later, the first three digits of the ICD-10 codes (ICD10CDX), along with three CCSR codes (CCSR1X, CCSR2X,

---

[4] Starting in 2018, data on informal telephone calls were no longer collected in the survey.

CCSR3X), are provided for each condition. Prior to 2016, MEPS conditions were coded into ICD-9-CM codes, which are generally less detailed than ICD-10 codes. Each ICD-9 code maps to a single Clinical Classification Software (CCS) code. For the 1996–2015 Medical Conditions files, the first three digits of the ICD-9 codes (ICD9CODX) along with the corresponding CCS code (CCCODEX) are provided.

Due to the different structure of the ICD-9 and ICD-10 codes, analyzing conditions before and after 2016 can have unexpected results. We recommend using extreme caution when conducting such analyses. When reporting results that cross the transition year, we recommend including clear caveats about the transition in figures and expository text.

Guidance #3: For treated priority conditions, limit analyses to 2008 and later.

Starting in 2007, a redesigned MEPS household interview was fielded, which included a major change in how priority conditions are collected in MEPS. The new design helps respondents and interviewers to link priority conditions to healthcare events more easily. This redesign resulted in a noticeable increase in the number of treated priority conditions reported in the medical conditions file (see Table 2). Apparent changes in healthcare use and expenditures related to priority conditions before and after the survey redesign may be an artifact of the changes to the survey, rather than actual differences in use and expenditures. Because 2007 was a transition year, we recommend limiting analyses of priority conditions to 2008 and later.

Table 2. Treated Prevalence by Priority Condition, 2005–2010

ICD9CODX	Priority Conditions	2005	2006	2007^	2008*	2009*	2010*
314	ADHD/ADD	3.9	4.0	4.3	5.1	5.5	5.7
413	Angina/Angina Pectoris	0.8	0.8	2.1	3.1	2.7	2.6
714-716	Arthritis	14.9	14.4	18.4	22.7	22.6	24.0
493	Asthma	12.1	13.2	13.1	14.8	15.0	16.1
140-239	Cancer	14.0	14.7	16.2	18.4	17.7	17.8
491	Chronic Bronchitis	0.4	0.3	2.0	4.3	3.5	3.5
414	Coronary Heart Disease	1.0	1.0	4.9	10.9	10.5	10.6
250	Diabetes/Sugar Diabetes	16.5	17.6	18.9	20.7	20.0	21.5
492	Emphysema	1.1	1.3	1.9	2.8	2.7	2.7
410	Heart Attack / Myocardial Infarction (MI)	1.5	1.5	2.9	5.0	5.2	5.2
272	High Cholesterol	28.6	30.2	36.9	44.5	45.2	46.1
401	Hypertension / High Blood Pressure	45.2	45.8	50.5	55.2	56.5	58.7
719	Joint Pain	8.0	8.2	10.9	15.3	15.9	17.0
436	Stroke / Transient Ischemic Attack (TIA) / Mini stroke	1.2	1.1	2.4	3.3	3.4	3.4

Note. ADD = attention deficit disorder; ADHD = attention-deficit/hyperactivity disorder. In this table, treated prevalence is defined as the number of people with at least one medical visit or prescribed medicine purchase associated with the condition. Medical visits include hospital stays, office-based or outpatient visits (including informal phone calls), ER visits, or home health care (including informal care).

^ 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 Condition Enumeration (PE) section was fielded to Panel 12.

* Cells with green shading indicate years in which the revised PE section was fielded.

Guidance #4: Avoid double-counting.

Some conditions in the Medical Conditions PUF may appear to be "duplicate" conditions at the ICD level. In the following example from the 2020 Medical Conditions PUF, the selected person has three condition records with ICD10CDX values of M54 ("Dorsalgia"). These are distinct records because the fully specified ICD-10 codes differ for these conditions (Table 3).

Table 3. Example of multiple condition records with the same ICD10CDX value

DUPERSID	CONDIDX	ICD10CDX	CCSR1X	CCSR2X	CCSR3X
2320529101	2320529101007	M54	MUS010	-1	-1
2320529101	2320529101008	M54	MUS011	MUS038	-1
2320529101	2320529101010	M54	MUS011	-1	-1

Because multiple conditions can be attributed to the same medical event or prescribed medicine purchase, it is important to de-duplicate when calculating healthcare utilization or expenditures at the ICD10CDX level. Note that the same guidance applies when analyzing conditions at the CCSR level because multiple conditions can map to the same CCSR value.

Guidance #5: Be mindful of the MEPS design when benchmarking.

Researchers are advised to use caution when comparing MEPS estimates to other surveys or sources. MEPS estimates may not benchmark to other sources for several reasons:

- **Survey population.** MEPS data represent the civilian noninstitutionalized U.S. population only. 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 better at reporting conditions that are highly salient, cause pain, need ongoing treatment or alter lifestyle, and/or affect daily life. Less salient conditions (e.g., the common cold, the flu) may be underreported due to recall error.

- **Proxy-reported conditions.** MEPS typically has one respondent who responds for every member of the household. The respondent may not have complete information for all people in the household, potentially resulting in underreporting or inaccurate reporting of medical conditions.

- **Condition accuracy.** Conditions reported in MEPS often lack the specificity and level of detail that would be common in a doctor's office or in administrative data. Reported conditions can be vague or ambiguous, either intentionally (e.g., respondent may not feel comfortable sharing details about a sensitive condition) or unintentionally (e.g., respondent lacks detailed information). Prior to 2020, interviewers would record reported conditions as verbatim text strings, which would later be coded into ICD-9 or ICD-10 codes. The addition of the condition pick-list tool in 2020 helped alleviate some of these issues by providing a searchable list of conditions that interviewers could use. This list allowed respondents to help clarify 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 respondents know and what they choose to report.