Medical Expenditure Panel Survey

MEDICAL CONDITIONS FILE

Medical Conditions File

Medical Conditions File contains

- **Household reported data**
- **Data on Conditions**
- **Data on Procedures**

This public use file provides information about the conditions for a nationally representative sample of the civilian noninstitutionalized population of the United States. All information on this file has been collected from household respondents only.

Each record represents one medical condition reported by a household survey respondent who resides in an eligible responding household and who has a positive person or family weight.

Although this file is called “Medical Conditions File”, it contains information on medical conditions and procedures.

Sources of Data

Sources of data on Medical Conditions File

- Condition Enumeration
- Medical Events
- Disability Section

Conditions can be added to the MEPS condition roster in one of several ways. Most directly, the condition can be identified as the reason reported by the household respondent for a particular medical event (hospital stay, outpatient visit, emergency room visit, home health episode, prescribed medication purchase, or medical provider visit). Second, the condition may have been reported as the reason for one or more episodes of disability days. Finally, the condition may have been reported by the household level respondent as a condition "bothering" the person during the reference period.

General File Structure

- **Each record represents a unique condition or procedure reported by a household respondent**
- **Depending on the number of conditions reported, persons may be represented on the file**
  - **once**
  - **several times**
  - **not at all**

Each record represents a unique condition (defined on the file by an ICD-9 condition code, V-code, or procedure code) reported by a household respondent. If a person reported 3 different conditions (e.g., asthma, hypertension, and diabetes) then she will have 3 separate records on the condition file. If another respondent reports asthma, bronchitis, and heart disease then he will be represented 3 times, with 3 separate records on the condition file. Even though each person reported asthma, there will be 2 separate records for asthma on the file -- one for each respondent who reported asthma.

If a person does not report any condition in a particular year, they will not have any records on the condition file.

If a person reports multiple episodes of an acute condition over the course of a year, multiple records will exist for that condition on the condition file.

Each record can include information about a condition or a procedure or both.

Reporting and Recording Conditions

- **Interviewer records verbatim text reported by the household respondent**
  - **Open-ended questions**

Example:

We're interested in learning about health problems that may have bothered (PERSON) {since (START DATE)/between (START DATE) and (END DATE)}. Health problems include physical conditions, accidents, or injuries that affect any part of the body as well as mental or emotional health conditions, such as feeling sad, blue, or anxious about something**.**

Respondents are asked to report current conditions in the condition enumeration section of the questionnaire at every round of data collection. This is an open-ended question and interviewers record the respondent's verbatim text. From this information CAPI generates a condition roster for every person in the household. Later in the interview, respondents are asked the reason for a medical visit, missed workdays, missed school days, and bed-days. At these points in the interview conditions can be added to the roster if they were not previously mentioned.

Reporting and Recording Conditions

- **Respondents may report having the same condition more than once**
  - **Interviewer verifies that these are different occurrences of the condition**
  - **each unique episode of a condition is recorded only once**
    - **person may have more than one cold in a year**
    - **each cold has a separate record**

If a respondent reports bronchitis in round 1 and again in round 2, the interviewer verifies whether or not this is the same condition that was reported in round 1. If it is a different condition then it is entered a second time. If the respondent indicates that it is the same condition reported in round 1 the interviewer does not enter a new condition. Similarly, if “bronchitis” is reported in round 1 and later in the interview given as the reason for a provider visit, the interviewer asks if this is the same condition reported previously. If it is a different condition, then it is entered on the condition roster again.

If a person reports multiple episodes of an acute condition over the course of a year, multiple records will exist for that condition on the condition file.

For chronic conditions, such as diabetes or hypertension, there will frequently be several provider visits reported for the same condition. But the condition will appear only once on the person’s condition roster and that person will only have one record for that condition on the condition file.

Priority Conditions

- **Designation based on**
  - **Prevalence**
  - **Expense**
  - **Relevance to policy**

In the Medical Conditions File, certain conditions are designated as “priority” conditions. See Appendix 4 of the file documentation for the complete list of priority conditions.

Priority Conditions

- **Long-Term, Life Threatening Conditions:**

  Cancer (of any body part)
  Diabetes
  Emphysema
  High Cholesterol
  HIV/AIDS
  Hypertension
  Ischemic Heart Disease
  Stroke prevalence

A complete list of these priority conditions is available in the Medical Conditions file documentation. During the interview, after a condition is mentioned, the interviewer looks at this “Priority Conditions” list to determine whether the condition is a priority condition or not. If it’s on this list, the interviewer checks it as such. This action generates a flag (PRIOLST = 1) for this record.

Priority Conditions

- **Chronic, Manageable Conditions:**
  - **Arthritis**
  - **Asthma**
  - **Gall Bladder Disease**
  - **Stomach Ulcers**
  - **Back Problems of Any Kind**
- **Mental Health Issues:**
  - **Alzheimer's Disease and Other Dementias**
  - **Depression and Anxiety Disorders**

These are some additional priority conditions.

Priority Conditions

- **Specific Priority Condition Questions**
  - **Date condition began**
  - **Round-specific questions**
    - Was a doctor ever seen or talked to for the condition?
  - **Follow-Up Questions in Later Round**
    - Still being treated for this condition?
    - How seriously does condition affect overall health and well-being?
    - Did a provider recommend further treatment?

If a condition record is flagged as PRIORLST = 1, the respondent is asked specific questions related to that condition. These questions start with asking “when did the condition begin”? Note that the ‘date condition began’ is not necessarily during the time the person is in the survey. The date a condition began is often many years prior to the survey interview.

Accidents and Injuries

- **Ascertained at time of interview**
  - **date of accident**
  - **place (work, home, school, etc.)**
  - **cause (gun, vehicle, fall, fire, etc.)**
  - **whether or not the person has recovered from the injury**

When a condition is first mentioned, respondents are asked whether it was due to an accident or an injury. If the condition was due to an accident or injury, date of accident, place of injury and cause of injury was asked. Lastly, the interviewer asks if the person has fully recovered from the injury.

Back problems – priority condition, if due to an accident data appears on the injury record.

Priority and Injury Conditions

- **Date Information**
  - If a condition is both a priority condition and an injury, date information is on the injury record.
- **Previous Year's Information**
  - Due to the overlapping panel design, some conditions may have been first reported in the previous year. Unless it was a Priority Condition or was linked to an event in 2004, it may not appear on 2004 data file.

Some priority conditions can also be identified as "injuries".

Only priority conditions and injuries have date information on the file.

If a condition is both an injury and a priority condition, only the injury date variables will have date information.

In Panel 9: All Round 1 and Round 2 conditions; Round 3 conditions that are linked to a 2004 event; Round 3 conditions that were on the priority list, not due to an accident or injury, and began before 2004; Round 3 conditions that were due to an accident or injury and began before 2004 Round 3 conditions where 50 percent or more of person's reference period occurred in 2004.

In Panel 8: All Round 4 and Round 5 conditions; Round 1, Round 2, and Round 3 conditions that meet at least one of the following two criteria: The condition is linked to a 2004 event; The condition is a priority condition; Round 3 conditions that are injuries; Round 3 conditions that were not previously delivered in the FY 2003 Conditions PUF (HC-078). This includes: Round 3 conditions created after the delivery of the FY 2003 Conditions File due to Round 4 and Round 5 comments processing; Round 3 conditions where the person did not have a positive person or family weight in FY 2003 but has a positive person or family weight in FY 2004; Round 3 conditions where fifty percent or more of person's reference period occurred in 2004.

Disability Flag Variables

- **Three Flag Variables**
  - **Missed work day (MISSWORK)**
  - **Missed school day (MISSSCHL)**
  - **Day spent in bed (INBEDFLG)**

Medical Conditions file contains three flag variables which indicate any disability days. They are coded 1 or 0 based on Yes or No answer to the questions. Due to the MEPS instrument design, there is no link indicating the specific number of disability days associated with a particular medical condition.

Condition and Procedure Coding

- **Coding and Editing**
  - **Fully specified ICD-9 CM codes (up to 5 digits)**
  - **ICD-9 condition codes collapsed to 3 digits to maintain confidentiality**
  - **Approximately 10% of condition codes are collapsed further by combining 2 or more 3-digit codes**
  - **Procedure codes collapsed from fully specified (up to 4 digits) to 2-digit codes**
  - **Approximately 3% of procedure codes are collapsed further by combining 2 or more 3-digit codes**

To preserve confidentiality, all of the condition codes were collapsed from fully-specified codes to 3-digit code categories. Table 1 in Appendix 2 of the file documentation provides unweighted and weighted frequencies for all ICD-9 condition code values. To further preserve confidentiality, approximately 10% of the ICD-9 codes were collapsed even further.

Procedure codes were also collapsed from fully specified codes to 2-digit category codes. To preserve confidentiality approximately 3% were further collapsed. Table 2 in Appendix 2 of the file documentation provides unweighted and weighted frequencies for procedures.

Procedures were under reported on the condition roster by household respondents as a reason for a provider visit. Analysts should use procedures identified on the event files for more accurate estimates of procedures.

Clinical Classification Codes

- **Formerly Clinical Classification for Health Policy Research (CCHPR)**
- **ICD-9 codes aggregated into clinically meaningful categories**
- **Edited to preserve confidentiality**

Clinical Classification (CCC) Codes were formerly known as CCHPR. CCC aggregates ICD-9 condition codes and V-codes into 260 mutually exclusive categories most of which are clinically homogeneous and clinically more meaningful. Table 3 in Appendix 2 of the file documentation provides unweighted and weighted estimates for the Clinical Classification Codes. Appendix 3 lists the ICD-9 codes that have been aggregated for each CCC category. Below is an example for diabetes.

49 Diabetes mellitus without complication

25000 25001 7902 79021 79022 79029 7915 7916 V4585 V5391 V6546

50 Diabetes mellitus with complications

25002 25003 25010 25011 25012 25013 25020 25021 25022 25023 25030 25031 25032 25033 25040 25041 25042 25043 25050 25051 25052 25053 25060 25061 25062 25063 25070 25071 25072 25073 25080 25081 25082 25083 25090 25091 25092 25093

Utilization Variables

Following Utilization Variables are included in the Medical Conditions file:

- **ERNUM**
- **HHNUM**
- **IPNUM**
- **OBNUM**
- **OPNUM**
- **RXNUM**

The variables ERNUM , HHNUM, IPNUM, OBNUM, OPNUM, and RXNUM indicate the total number of events that can be linked to each condition record on the current file, i.e., emergency room visits, home health, inpatient hospital stays, office-based, outpatient, and prescribed medicines, respectively.

These counts of events are derived from Expenditure Event Public Use Files. Events associated with conditions include all utilization that occurred between January 1 and December 31.

Utilization Variables

- **Events for Multiple Conditions**
  - **Events may be associated with more than one condition**
  - **Example: One hospital stay for 3 conditions**
    - **Fractured hip, fractured shoulder, concussion**

Because persons can be seen for more than one condition per visit, the condition level frequencies will not match the person or event-level utilization counts. For example, if a person had one inpatient hospital stay and was treated for a fractured hip, a fractured shoulder and a concussion, each of these conditions has a unique record in this file and IPNUM=1 for each record. By summing IPNUM for these records, the total inpatient hospital stays would be three when actually there was only one inpatient hospital stay for that person and three conditions were treated. These variables are useful for determining the number of inpatient hospital stays for head injuries, hip fractures, etc.

Weighted Estimates

- **Medical Conditions file includes person-level weights**
  - Frequencies from this file will estimate the number of times a condition was reported by the sample population
  - Number of persons reporting a condition can only be estimated at the person level

Frequencies derived from the condition-level file will estimate the number of times a condition was reported, e.g. you may want to find the number of head injuries reported in a particular year. If you want to find the number of persons who reported diabetes or asthma you will need to do this at the person level.

Limitations of Condition Data

- **Limitations**
  - inaccurate or vague reports of condition
  - clustering of ICD-9 codes in NEC (not elsewhere classified)
  - one respondent provides information for the entire household

The medical condition coding was done by qualified and experienced medical coders. The error rates were less than 2.5% for condition coding. In spite of these measures, analysts are cautioned about the limitations of the medical condition data. The ability of household respondents to report condition data that can be coded accurately should not be assumed. Respondents are sometimes vague about the conditions, which results in clustering of ICD-9 codes as “Not Elsewhere Classified”. In addition, one household respondent reports this information for all members of the household.

Because of these limitations, it’s not recommend that analysts use condition data for prevalence or mortality studies.

Merging/Linking to Other MEPS Files

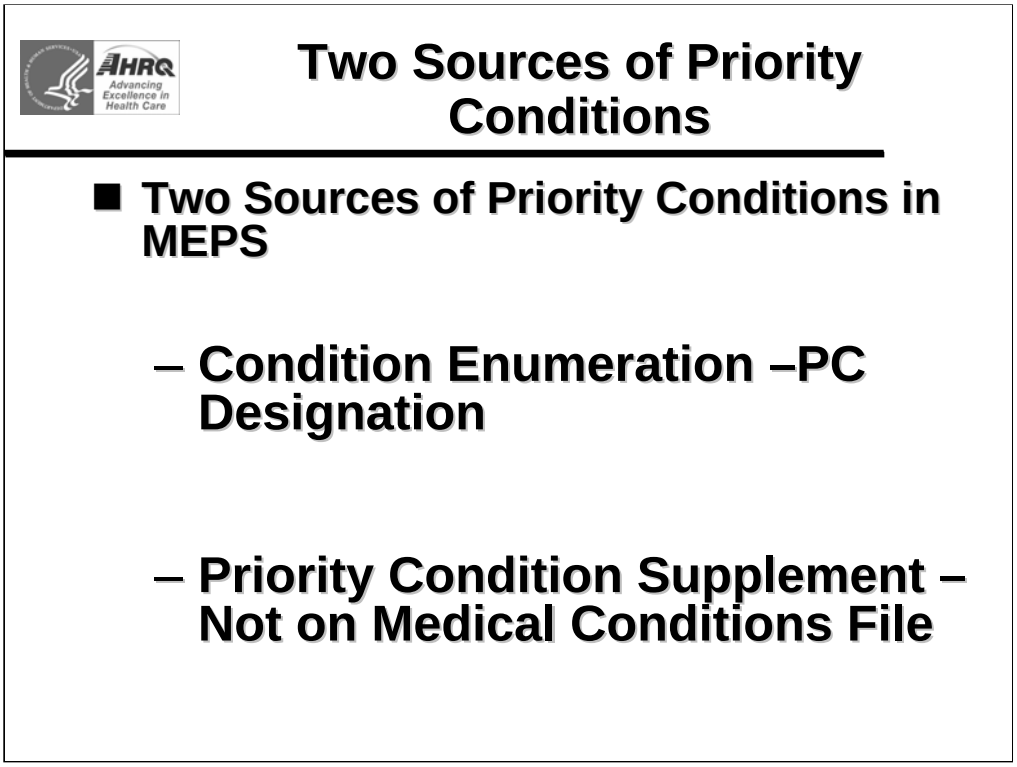
- **ID Variables**
  - used to
    - identify and distinguish records on a file
    - match records in different files
  - **DUPERSID** (person-level ID)
  - **CONDIDX** (condition-level ID)
  - **EVNTIDX** (event-level ID)
    - but note PMED differences (LINKIDX, RXRECIDX)
  - **CLNKIDX** (condition-event link ID)

The MEPS PUF files are constructed so that each record (or, row) contains ID variables which allow you to associate the information on that row with a person or an event.

Each file has a *key* variable which uniquely identifies a row. For example, the full-year consolidated data file is a *person-level* file, i.e. each row has data for one person and different rows represent different persons. The key variable on this file is DUPERSID.

The medical conditions file is a *person-condition-level* file. The key variable CONDIDX is composed of DUPERSID plus a condition number. Each CONDIDX (each row on the condition file) uniquely identifies a condition for a person.

ID variables are also important for merging different files. Merging files allows you to make connections between different files, e.g. between the medical conditions file and the hospital stays event file– or between the conditions file and the PMED file.



The other source of data on priority conditions is the Priority Conditions (PC) Supplement. AHRQ is mandated by Congress to provide a National Healthcare Quality Report. MEPS added a battery of questions for this purpose. These questions were added to identify persons with certain priority conditions. Unlike other MEPS condition data that is conditioned on the reference year, this information is asked in the framework of "Did a doctor or health professional ever tell you that you had (CONDITION)?" The conditions enumerated in this section are not added to the condition roster. These conditions are reviewed periodically and will be subject to future enhancements.

No expenditure information is collected for the PC section. This is a measure of numbers of individuals who were ever told by a health professional that they had a certain condition. The numbers you will get from this analysis will be different than the numbers in the conditions file for the same condition (e.g., Diabetes).