

# 2017 Medicare Current Beneficiary Survey (MCBS): Frequently Asked Questions



## Version Control Log

Date	Version	Revisions
10/16/2019	1.0	Initial version published.

# TABLE OF CONTENTS

<b>Frequently Asked Questions</b> .....	<b>1</b>
Data Requests.....	1
File Corrections .....	1
Content .....	2
Sampling .....	5
Analysis .....	5

## FREQUENTLY ASKED QUESTIONS

### Data Requests

- **How do I request the MCBS Limited Data Set (LDS) files and how long does it take to receive the data?** Information on how to request the MCBS LDS files is posted on the CMS website linked here: <https://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/LimitedDataSets/index.html>

Please also see the instructions on the [DUA - Limited Data Sets](https://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/Data-Disclosures-Data-Agreements/DUA_-_NewLDS.html) page for more information on the request process: [https://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/Data-Disclosures-Data-Agreements/DUA - NewLDS.html](https://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/Data-Disclosures-Data-Agreements/DUA_-_NewLDS.html)

The processing of DUAs takes approximately 6-8 weeks. If approved, then data processing time varies depending on the number of data years and files requested. Normal data processing time is one week.

The MCBS Public Use File (PUF) is available to the public as a free download and can be found through CMS' PUF website at <https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/MCBS-Public-Use-File/index.html>.

### File Corrections

- Occasionally, there are corrections to the original files that a data user has received. Recipients will be notified via email, and, where possible, a zipped/encrypted correction will be sent via email.
  - ▶ **If you did not receive the updated files.** If you received a password notification but did not receive updated MCBS files, it is likely that your organization has removed the attachments (.zip files). Check your e-mail security settings to determine how attachments are handled by your organization. The updates are also available as executable (.exe) files. If executable files are supported by your organization, these can be sent to you by request.
  - ▶ **If the "password" doesn't work.** MCBS data products are encrypted using PKWARE and this can sometimes cause problems when opened with other encryption packages. If the error you are receiving is asking for a password, this is the problem. PKWARE uses the term "passphrase" instead of password. PKWARE is freeware, if needed. For more information, please see:  

SecureZIP® by PKWARE, Inc.  
<https://support.pkware.com/>
  - ▶ **If you received multiple emails from the MCBS.** Updates are sent by DUA#. If you receive password notifications or receive data more than once, it is likely because you are listed as having multiple DUAs that include MCBS data.

## Content

- **How do the 2015-2017 Health Insurance Timeline (HITLINE) and Health Insurance Summary (HISUMRY) segments differ from the previously released segments (i.e., RICs 4 and A)?**
  - ▶ File Structure:
    - HITLINE is a Beneficiary-Plan level file, which includes the health insurance coverage timeline for administrative and survey-only plans for Community and Facility beneficiaries. It also includes detailed information for survey-reported plans. HISUMRY is a Beneficiary-level file for Community and Facility beneficiaries, which includes detailed information for administrative plans.
    - These files differ from the previously-released versions of the files in 3 major ways:
      1. In older versions of HITLINE, some plans appeared more than once. This was due to a processing/reporting issue that led to some plans reported over several rounds not being collapsed into a single plan. This issue has been corrected.
      2. Previously, plan details for up to 5 survey-reported plans appeared in HISUMRY. Now, plan details appear in HITLINE for all survey-reported plans.
      3. HISUMRY now includes information for administrative plans *only*. The only exception is the EST\_TPRM variable, which is the estimated sum of annual premium amounts for all plans (Part A + Part B + Part C + Part D + Survey Only Plan premiums).
        - a. Part A premiums: This is the dollar amount from beneficiaries who purchased Part-A by paying a monthly premium (a relatively small number of beneficiaries).
        - b. Part B premiums: This is the dollar amount from all Part-B beneficiaries (a large number; a premium is always paid, by either the beneficiary or a third party). NOTE: There will be no Part-B premium paid indicated if the beneficiary belongs to a managed-care plan in which the plan pays the entirety of the premium. Since the managed-care plan paid the entire premium, the data show no premium paid.
        - c. Part C premiums: This is the dollar amount of the Medicare Advantage (referred to as Medicare Part C) Basic Plus Mandatory Supplemental Premium Rate (Net of Rebates). The Part C premium for Medicare Advantage Plans, Cost Plans, and Demonstrations covers Medicare medical and hospital benefits, and supplemental benefits, where offered. Beneficiaries generally are also responsible for the Part B premium. The data source is constructed from information submitted by Part D plan sponsors to CMS's Health Plan Management System (HPMS). Employer and National Program of All-inclusive Care for the Elderly (PACE) plans are waived from reporting Plan Benefit Package information. For those plans that did not report, the value of this variable will be blank.
        - d. Part D premiums: This is the dollar amount of the Part D Total Premium (basic + supplemental) Rate (Net of Rebates). The Part D Total Premium is the sum of the Basic and Supplemental Premiums. This amount is net of any Part A/B

rebates applied to "buy down" the drug premium for Medicare Advantage plans; for some plans, the total premium may be lower than the sum of the basic and supplemental premiums due to negative basic or supplemental premiums. The data source is constructed from information submitted by Part D plan sponsors to CMS's Health Plan Management System (HPMS). Employer and National Program of All-inclusive Care for the Elderly (PACE) plans are waived from reporting Plan Benefit Package information. For those plans that did not report, the value of this variable will be blank.

- e. Survey Only Plan premiums: This information is sourced from the D\_ANNPRM variable in the HITLINE file. D\_ANNPRM is the annualized premium that the beneficiary pays for their health insurance plan. It is obtained from the survey and reflects information for survey-reported plans ONLY. D\_ANNPRM is set to missing for all administrative plans. Premium information associated with a plan would not be included in D\_ANNPRM (and, hence, EST\_TPRM) unless a beneficiary reported both the premium amount paid and the frequency of those payments. Data users interested in making different assumptions about how premium information for survey-reported plans should be handled can utilize information provided in the HITLINE segment, as it contains all available data for each survey-reported plan. Note that because MCBS Facility interviews are conducted with facility staff rather than beneficiaries, the survey has no information on premiums paid by these beneficiaries for survey-only plans. For beneficiaries residing in a facility setting, EST\_TPRM reflects Part A – D premiums only.

- Private plans reported in a community setting can be linked across years using the unique plan identifier, PLANNUM.
  - HITLINE files now indicate why certain information is missing for a given plan. The value of 'N' (Not ascertained) refers to instances where a beneficiary was supposed to receive a question but did not, which can occur for several reasons. The value of '.' (Inapplicable/Missing) refers to instances where a beneficiary was out of universe for a given question (for example, all beneficiaries who report not paying a premium will have a value of '.' for the premium amount associated with that plan).
- ▶ Community vs. Facility Plans:
- Plans reported in a facility setting have fewer details about coverage collected, which is a consequence of data collected from facility staff, rather than the beneficiary. The HITLINE files include a flag, D\_FCLTYF, which makes it possible to identify plans that are missing certain plan details because they were reported in a facility setting.
  - Given that only limited information about coverage type is collected from facility staff, it is not possible to determine whether a plan reported in a facility is, in fact, the same plan that was also reported in a community setting and / or a plan reported in a facility setting at a different point in time.
- ▶ File-Specific Notes:

- PLANTYPE can change from year to year for some plans. This is a consequence of beneficiaries (or a proxy) providing slightly different information about a plan, which can result in changing classification from year to year. While responses are broadly consistent over time regarding whether a plan provides comprehensive coverage, there is less consistency regarding whether the plan was employment-based or purchased directly.
- **What type of Medicare eligibility/enrollment data are included?** The Survey File LDS contains information on Medicare eligibility and enrollment data. Specifically, the HITLINE segment provides monthly coverage indicators, coverage start and end dates, the type of plan, and the source of coverage information for the plan. The HISUMRY segment also contains eligibility codes and detailed Medicare-Medicaid dual eligibility indicators.
- **Are Medicare Advantage (MA)/Medicaid/Part D claims data included?** The Survey File release contains the Fee-for-Service (FFS) claims data, which provide CMS administrative information on medical services and payments paid by Medicare FFS claims; claims data for Medicare Advantage beneficiaries are not available, nor are Medicaid claims or Part D claims. However, the MCBS data does include survey-reported healthcare utilization and cost data for MA beneficiaries in the Cost Supplement.
- **What cost and utilization information is available for beneficiaries enrolled in Medicare Advantage (MA)?** When a beneficiary reports healthcare events, we use the explanation of benefits (EOBs) form from their Medicare Advantage provider to report the payments, as well as the capitation information from the administrative data for total Medicare Advantage Payments. This is the same approach we take for services that are not covered by Medicare, such as most dental care. Actual claims-based information for MA beneficiaries, referred to as encounter data, are not currently available for these individual events.
- **How often do respondents receive each questionnaire?** Different combinations of MCBS Questionnaire sections are used depending on a number of criteria, including interview type (Baseline vs. Continuing); the season of the round of data collection (fall, winter, summer); whether the respondent is alive, deceased, or in a facility; and whether the interview is being completed with the beneficiary or a proxy. For more information about the specific questionnaires administered during each round of data collection, please see the *MCBS Data User's Guide: Survey File* available on CMS' MCBS website. Questionnaires for each data year are also available on CMS's MCBS website at <https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS/Questionnaires.html>.
- **How is the Cost Supplement ever-enrolled population derived?** To define the population, the MCBS creates a calendar history of a beneficiary's MCBS interviews. A number of eligibility checks are run against this calendar history to identify beneficiaries who met eligibility requirements for inclusion in the survey data for the calendar year, either because they were interviewed for a full year or interviewed until death or loss of Medicare entitlement. Beneficiaries who pass these eligibility checks become the population eligible for the Cost Supplement ever-enrolled weight and the prescription medicine data files.

## Sampling

- What types of beneficiaries are in the continuously enrolled vs. ever enrolled populations?** The continuously enrolled represent a population of beneficiaries who were enrolled continuously between January 1st of the data year and the completion of their fall interview. Beneficiaries who died during the calendar year, newly-enrolled beneficiaries who enrolled in Medicare during the year that they were sampled, and beneficiaries who have lost eligibility are not included in the continuously enrolled group. The ever enrolled represent the population of beneficiaries who were ever enrolled in Medicare for at least one day at any time during the data year. The ever enrolled population includes beneficiaries who died or lost entitlement prior to completing the fall interview. Beneficiaries who first became enrolled in Medicare during the data year are also included. Thus, the continuously enrolled beneficiaries are a subset of the ever enrolled beneficiaries.
- Does the survey use a household sample or a list sample?** The survey uses a list sample. The sample for the MCBS is drawn from a subset of the Medicare enrollment data, which is a list of all Medicare beneficiaries.
- Do Primary Sampling Units (PSUs) and Secondary Sampling Units (SSUs) align with other federal health surveys, such as NHIS?** The MCBS selects its own PSUs and SSUs. In late 2000, the current set of PSUs was selected. In 2014, SSUs were reconstructed using Census tracts, and a new sample was drawn. While the MCBS PSUs and SSUs do not align directly with other surveys, they may overlap in some areas with PSUs and/or SSUs used for other federal health surveys.
- Are populations (given changes to the sample design, e.g. the addition of newly-eligible beneficiaries to the sample) comparable with past years?** The Survey File cross-sectional and longitudinal population definitions are consistent from year to year, so the data are comparable between years. The Cost Supplement cross-sectional population definition is also consistent and comparable from year to year. The Cost Supplement two-year longitudinal population changed slightly in 2016 from what was defined the last time the two-year longitudinal weights were supplied (i.e., in 2013). In 2013, the two-year longitudinal (i.e., one-year backward longitudinal weight) Cost Supplement weights represented the population that enrolled on or before 1/1/2011 and was still enrolled in 2013 (i.e., enrollees after 1/1/2011 were not included). Beginning in 2016, the two-year longitudinal weights represent a true two-year ever-enrolled population (i.e., the population of beneficiaries who were ever enrolled in both 2015 and 2016).

## Analysis

- Can the MCBS be used to produce subnational estimates? / Are MCBS estimates representative at the state or local level?** The MCBS is designed to produce nationally representative estimates of the population of all Medicare beneficiaries. Estimates from MCBS data are not representative at the state or local level.
- Why do I see differences between Medicare published statistics and estimates using MCBS data?** In general, MCBS estimates may differ from Medicare program statistics using 100 percent administrative enrollment data. There are several reasons for the differences. The most important reason for the difference is that the administrative enrollment data may include people

who are no longer alive. This may occur where people have entitlement, such as for Part A only, and receive no Social Security check. When field interviewers try to locate these beneficiaries for interviews, they establish the fact of these deaths. Unrecorded deaths may still be present on the Medicare Administrative enrollment data. The MCBS makes every effort to reconcile the survey information against the administrative data when possible. Other reasons, such as sampling error, may also contribute to differences between MCBS estimates and Medicare program statistics. Lastly, estimates may differ because Medicare program statistics adjust for partial enrollment. Medicare program estimates use a 'person year' calculation where partial enrollment is counted as a fraction for the year. In contrast, the MCBS gives each beneficiary the same weight regardless of full or partial enrollment during the year, thus leading to differences in estimates using Medicare published statistics and MCBS data.

- **Where can I find documentation for longitudinal analysis?** *The Data User's Guide: Survey File* contains information on conducting longitudinal analyses (see section 9). *The Data User's Guide: Cost Supplement File* also contains some sample analyses with programs (see Technical Appendix C.1).
- **Why does the number of BASEIDs differ across segments?** There are multiple reasons why the number of BASEIDS may differ across segments. First, some segments include data from Community questionnaires and others from Facility questionnaires with different numbers of beneficiaries providing responses. Second, there are also differences in the number of beneficiaries by the specific round completed. Third, the use of ever enrolled or continuously enrolled weights in constructing the segments may result in differences.
- **How can I conduct subgroup analyses and maintain the appropriate variance estimation?**

Variance estimation can be impacted by selecting individuals prior to analysis. If the BRR variance estimation method is used, subgroup analyses can be conducted by limiting the dataset to the desired sub-sample. If other variance estimation methods, such as Taylor Series linearization are used, the correct way to analyze MCBS data is to employ domain statements (procs `surveymeans`, `surveylogistic`, and `surveyreg`) or indicator variables in three-way tables (proc `surveyfreq`).

For indicator variables in three-way tables, you can create flags to help you identify the population of interest. The variables SUDSTRAT (sampling strata) and SUDUNIT (primary sampling unit) are included for variance estimation using the Taylor Series linearization method. This method does not require replicate weights. For instance, if you are interested in the prevalence of diabetes in men versus women, but only in the over-65 population in Medicare Advantage, you could use the following SAS code:

```
proc surveyfreq data=mcbsdata;

    table GENDER * DIABETES * FLAG / col notot;
    strata SUDSTRAT;
    cluster SUDUNIT;
    weight CEYRSWGT;

run;
```

...where the FLAG variable is set to 1 if the beneficiary is over 65 and in Medicare Advantage, 0 otherwise (for example).

Since variance estimation using the BRR approach permits limiting the dataset to the desired sub-sample of interest, the following SAS code can also be used to achieve the same result.

```
data mcbsdata_subset;

    set mcbsdata;
    if FLAG = 1 then output;

run;

proc surveyfreq data=mcbsdata_subset VARMETHOD = brr (fay=.30);

    table GENDER * DIABETES / col notot;
    weight CEYRSWGT;
    repweight CEYRS001 - CEYRS100;

run;
```

- **Can I use the Survey File longitudinal weights with the Cost Supplement data?** The Survey File longitudinal weights are for analysis of Survey File data. Data users cannot use the Survey File longitudinal weights with Cost Supplement data. There are no longitudinal weights for the 2015 Cost Supplement, because 2014 data were not released. Beginning with the 2016 data year, there is a two-year longitudinal Cost Supplement weight. In the 2017 data year, there are two- and three-year longitudinal Cost Supplement weights. Users who want to analyze Survey File data along with utilization and cost data in the Cost Supplement should limit analysis to cases with a positive Cost Supplement weight.
- **Which cross-sectional weights are available in which data years?** Cross sectional weights are available for the Survey File and the Cost Supplement File in each data year. The Survey File LDS contains weights for the continuously enrolled (CENWGTS) and ever enrolled (EVRWGTS) populations. Given that the Cost Supplement population represents an ever enrolled population enrolled in Medicare on at least one day at any time in 2017, the Cost Supplement LDS contains cross-sectional weights for the ever enrolled population only (CSEVWGTS). The population represented by the sum of CSEVRWGT is identical to the population represented by the sum of the ever enrolled Survey File weight, but it is populated for a smaller subset of respondents with complete cost and utilization data.
- **When I use the topical segment weights, I see higher numbers of beneficiaries when using the continuously enrolled weight vs. the ever enrolled weight. Why is this the case?** The weights for the topical segments only include weights for the Survey File continuously-enrolled population and the Cost Supplement File ever-enrolled population. There are no weights included for the Survey File ever-enrolled population. As such, the number of beneficiaries with a continuously-enrolled weight is much greater than the number of beneficiaries with an ever-enrolled weight in all of the topical data segments. This means that you can use the Cost

Supplement ever-enrolled weights to analyze the topical content in conjunction with the Cost Supplement data.

- **If I am using the variables from the 2016 Survey File DENTAL segment DVNEED and DVNDRS/DVNDRS, is it more appropriate to use the Survey File weights or Cost Supplement weights for analysis?** When conducting data analysis with the variables DVNEED and DVNDRS/DVNDRS, it is most appropriate to use the Cost Supplement weights (i.e., CSEVRWGT). The reason for this is that the Cost Supplement population includes the full universe of persons who couldn't get dental care, although some missingness will be present.

The questions are asked in every round but in the following manner:

1. In non-fall rounds, respondents who are alive and reside in the community (that is only respondents who are in the community in that round) are asked these questions whether they report a dental event or not (so, may include people who had a dental event).
2. In fall rounds, only members of the Continuing panels that reside in the community who reported no dental utilization are asked these questions. No one from the Incoming Panels are asked the questions.

In summary, the universe of the variable DVNEED and DVNDRS/DVNDRS are all Continuing cases who are alive and not institutionalized, regardless of whether they reported a dental event. Facility cases need to be excluded and the variable SPLIVE must equal 1. The Incoming Panel will also be out of universe and can be excluded using the PANEL variable on the DEMO segment. Any remaining difference may be due to persons who become eligible for Medicare in the current year.

NOTE: In 2017, there is no longer a DENTAL segment. Instead, questions previously released in the DENTAL segment have been moved to the ACCSSMED segment with other questions regarding foregone healthcare and should use the topical weights included with that segment.

- **Can I link MCBS to electronic medical records?** MCBS data cannot be linked to electronic medical records, or to any other records that record lab values or physiologic data. MCBS data can be linked to Medicare Part A and Part B claims data for beneficiaries who participated in the MCBS.
- **Are physical exams performed for the MCBS?** No physical exams are administered as part of the MCBS. However the MCBS will implement new physical measures with additions to the Survey File LDS expected in 2020.
- **If I am using the variables from the 2013 and earlier data files for the Low-Income Subsidy (LIS) indicator H\_PDLS01-12, how do they compare to the 2015 and subsequent Survey File HISUMRY segment's new values?** When conducting data analysis with the variables H\_PDLS01-12 from the 2013 and earlier files and the 2015 and later files, you will need to recode the 2015 and later data to the previous values.

Before 2015, the LIS indicator had the following possible values:

- ▶ 1 - No premium subsidy
- ▶ 2 - 25% premium subsidy
- ▶ 3 - 50% premium subsidy
- ▶ 4 - 75% premium subsidy

- ▶ 5 - 100% premium subsidy

Beginning in 2015, the possible values changed to the following:

- ▶ 00 - Not Medicare enrolled for the month
- ▶ 01 - Pt A and/or B +D;elig LIS 100%,no copay
- ▶ 02 - Pt A and/or B +D;elig LIS 100%,low copay
- ▶ 03 - Pt A and/or B +D;elig LIS 100%,hi copay
- ▶ 04 - Pt A and/or B +D;enrld LIS 100%,hi copay
- ▶ 05 - Pt A and/or B +D;enrld LIS 100%,15% copy
- ▶ 06 - Pt A and/or B +D;enrld LIS 75%,15% copay
- ▶ 07 - Pt A and/or B +D;enrld LIS 50%,15% copay
- ▶ 08 - Pt A and/or B +D;enrld LIS 25%,15% copay
- ▶ 09 - Pt A and/or B +D;no prem or cstshare sub
- ▶ 10 - Pt A and/or B,no D;employr recvs RDS sub
- ▶ 13 - Pt A and/or B,no D;none of above cond

To compare the 2015 and later H\_PDLS01-12 values, they will need to be mapped back to the 2013 and earlier values as follows:

Exhibit 1: Mapping of H\_PDLS01-12 Values from 2013 and Prior to 2015 and Beyond

2015 and later H_PDLS01-12 values	Mapping to 2013 and earlier values
01,02,03,04,05	05
06	04
07	03
08	02
09,10,13	01

Note: The focus on the percent premium subsidy and the new data source for this variable (2015 and later) gives added detail on enrolled vs. eligible.