2015 MEDICARE CURRENT BENEFICIARY SURVEY (MCBS) – 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 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

  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.

Content

- **What type of Medicare eligibility/enrollment data is included?** The Survey File LDS contains information on Medicare eligibility and enrollment data. Specifically, the Health Insurance Timeline (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 Health Insurance Summary (HISUMRY) segment also contains eligibility codes, information about private plans, 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.
- **Does administrative data override survey-reported data?** In linking survey-reported and administrative data, we keep records from both sources to provide more complete data. Indicators in the file will usually tell you if the information is survey-reported only, administrative data only, or both. Data that are only administrative are indicated as such in the data documentation and codebook. In the HITLINE segment, administrative data does override survey-reported data if there is a discrepancy. For instance, if the respondent reported that the beneficiary is covered by Medicaid from January through December of the calendar year, but the administrative data shows coverage only for January through June, then the ENDDATE of coverage will be set to June 30. The Coverage source variables will indicate the differences. In this example, the SRCCOV01 – SRCCOV06 variables will have a value of 3 (Both Survey and Admin), but the variables SRCCOV07 – SRCCOV12 will have a value of 1 (Survey data only). The variables COV01 – COV06 will have a value of 1 (Eligible) and COV07 – COV12 will have a value of 0 (Ineligible).
- **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: General Information available on CMS' MCBS website.
- **How was race/ethnicity data collected in the MCBS**? Beginning in 2013, the MCBS began asking the extended Race/Ethnicity questions as required by OMB and HHS for all Federal surveys. Note that we do not have race/ethnicity as sampling domains as the information we get on the initial sample is very limited, but once we have the individual in the survey we ask race/ethnicity information in the demographic section of the questionnaire.

  Responses to race and ethnicity questions were recorded as interpreted by the respondent. Respondents who reported they were white and not of Hispanic origin were coded as white non-Hispanic; those who reported they were black/African-American and not of Hispanic origin were coded as black non-Hispanic; persons who reported they were Hispanic, Latino/Latina, or of Spanish origin, regardless of their race, were coded as Hispanic; persons who reported they were American Indian or Alaska Native, Asian, Native Hawaiian or other Pacific Islander, or other race and not of Hispanic origin were coded as other race/ethnicity. Respondents with more than one racial background were captured in a separate category and collapsed into the "other" category. In 2015, we included an oversample of Hispanics in the continental U.S. as previously the majority of our Hispanic sample came from Puerto Rico.

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 in 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 that died or lost entitlement prior to completing the Fall 2015 (Round 73) interview. Beneficiaries who first became enrolled in 2015 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 surveys, such as NHIS?**

  The MCBS selected 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 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 were still enrolled in 2013 (i.e., enrollees after 1/1/2011 were not included). Beginning in 2016, the two-year longitudinal weights now represent a true two-year ever-enrolled population (i.e., the population of beneficiaries that were ever enrolled in both 2015 and 2016).

Analysis

- **What is the smallest unit of analysis?** Although the MCBS data is nationally representative, it is not representative at the regional or state level so you cannot use the data to produce regional or state-level estimates. However, you can use the data to look for national trends across demographics.
- **What data are imputed?** MCBS imputation falls under two umbrellas that focus on imputing monetary amounts: Income and Asset (IA) imputation, and Event, Payer, and Cost imputation, which includes imputation for Prescription Medicine (PM) and Non Prescription Medicine (Non PM) events and costs. All three imputations focus on imputing a monetary amount. IA imputation completes income and asset information for the beneficiary and spouse, and PM and Non PM imputation complete medical event and cost data. For all three types, two groups of variables are imputed:
  - Probes: Yes/no variables indicating whether the type of income, asset, or payer should have a nonzero amount.
  - Amounts: The value of the income, asset, or cost paid for a medical event. For IA imputation, amounts are nonzero if the associated probe indicates the income or asset exists and missing otherwise. For PM and Non PM imputation, amounts are nonzero if the associated probe indicates that the payer paid and zero otherwise.

  For both probes and amounts, single value imputation is performed sequentially from variables or records with the least to the most item nonresponse. For more information on imputation, please see the MCBS Methodology Report and the Data User's Guide: Cost Supplement File (Technical Appendix C.2).

- **How does the claims match process occur? Does the data include claims for non-survey reported events, and vice versa?** The Survey File LDS contains the Fee-for-Service (FFS) claims data, which provide CMS administrative information on medical services and payments paid by Medicare FFS claims. In the Cost Supplement File LDS, survey-reported data are matched with administrative bill data to adjust for survey under-reporting using more complete administrative bill data, and to fill in and correct survey-reported payment amounts with more accurate information from bills submitted to and paid by Medicare. Note that this matching is only conducted for FFS beneficiaries accessing services covered by Medicare. The matching is not a process to drop records not matched. We keep records from both administrative and survey-reported sources. The process is more akin to a data linking process where we keep data from both sources even if they cannot be matched. A full description of the claims match process can be found in the Data User's Guide: Cost Supplement File (Technical Appendix C.2).

- **Where can I find documentation for longitudinal analysis?** The Data User's Guide: General Information contains information on conducting longitudinal analyses (see section 8.4.1). The Data User's Guide: Cost Supplement File also contains some sample analyses with programs (see Technical Appendix C.1).
- **What are the analytic limitations of using MCBS data?** Data users should consider the following points when conducting and designing analyses.
  - Potentially inadequate sample size to analyze some subgroups – Analysis of subgroups, such as decedents, new enrollees, or beneficiaries with selected medical conditions may not be feasible. Pooling data across years may mitigate this issue (see the Data User's Guide: General Information for further guidance on conducting subgroup analysis or pooling data across years).
  - Producing sub-national estimates may not be feasible – Since MCBS is a nationally representative survey, producing direct sub-national estimates is not appropriate. Advanced statistical techniques such as small area estimation may provide a way to bridge this.
  - Users should be cautious about comparing cost and utilization across MA and FFS – Since survey reported utilization is not linked to MA claims – MA encounter data is currently unavailable – cost and utilization associated with MA enrollees will not be as complete and accurate as FFS.
  - Users should be cautious about comparing Medicare beneficiaries in different settings – Since the survey instruments and data collection approaches are different for community and facility dwelling respondents, users should be cautious about drawing inferences about differences between these two populations.
- **Why do 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.
- **Why do I see differences in residence status (i.e., INT_TYPE variable on the DEMO segment) across segments?** In each data year, some differences by segment will exist with INT_TYPE, the variable on the DEMO segment that provides the beneficiary's residence status at the time of interview (C=Community, F=Facility, or B=Both). INT_TYPE is calculated on the benefit year, but data segments may reflect a prior or future calendar year due to the specific questionnaire and reference period used to collect the information. That is, the segment data is collected prior to or after the benefit year designation of INT_TYPE.

  For example, there are two to four Facility-dwelling beneficiaries (INT_TYPE = F) that appear on the 2015 segments that include 2016 non-response adjustments: INCASSET, FOODINS, MCREPLNQ, RXPARTD, PNTACT. We would expect these segments to only include beneficiaries with INT_TYPE = C or B because these segments contain data from survey-reported instruments only asked of beneficiaries that reside in the community. However, due to the fact that the data for these segments is collected in 2016, beneficiaries may have moved from a facility in 2015 to the community in 2016 at the time these data segments were collected.

  Alternatively, data may be pulled forward from a prior data collection year. For example, a beneficiary in 2013 that answered affirmative to the question, "Have you ever had a hysterectomy?", a survey item that is asked of beneficiaries in the Fall in the Community questionnaire, will have that answer pulled forward to

the 2015 data segment even if the beneficiary currently resides in a facility in 2015, and thus they would show an INT_TYPE = F.

INT_TYPE is only constructed using survey-reported data for the benefit year and is not edited to account for data collected in a future or prior data year. For more information on the construction of INT_TYPE, please see the DEMO segment FAQs below.

- **Were the MAFLAG variables derived from the H_ENT and H_PLTP variables?** The MAFLAG variables are the most reliable indicators for monthly MA information, as of the 2009 MCBS files. This information is sourced from the CMS administrative data. The MAFLAG variables are found on the HISUMRY segment of the Survey File LDS (H_MAFF01-12). The H_PLTP variables were removed from the RIC A file after 2008.
  The H_ENT variables were used to determine if the individual did not have Medicare entitlement. This information factored into the "No Entitlement" category in the MAFLAG monthly variables. The monthly entitlement variables can be found on the HITLINE segment. H_DOE and H_DOT on the HISUMRY file provide Medicare entitlement start and end dates for the beneficiary.
  Because the administrative source of this information has changed, H_ENT variables cannot be used to "crosswalk" to the MAFLAG variables. H_ENT can be used to determine Part A and Part B eligibility among FFS beneficiaries in files prior to 2015.
  Regarding the PLTP variables: In the past we released these variables as part of MCBS data. While there is a distinction between Medicare Advantage (MA) and Fee-for-Service (FFS), questions arose regarding the more detailed definitions or categorizations of MA. The MAFLAG variables (displaying "FFS" vs. "MA") were created as a replacement for the PLTP's lower level definitions. The decision was made to use the MAFLAG variables as our data was better aligned to support that level of granularity across the years.

- **What happened to VARSTRAT and VARUNIT?** VARSTRAT and VARUNIT have been discontinued. These variables had no utility except for the development of the replicate weights, which most of our users do not compute independently since we provide them in the data release. If a researcher does need VARSTRAT and VARUNIT for this purpose, we can provide them on an as-needed basis.
  The MCBS includes variables to obtain weighted estimates and estimated standard errors using either the Taylor-series linearization approach or balanced repeated replication. These variables, include the cross sectional weight (CS1YRWGT), the replicate weights (CS1YR001- CS1YR100), the sampling strata (SUDSTRAT), and the primary sampling unit (SUDUNIT).
  To compute the standard errors with MCBS replicate weights, a form of the balanced repeated replication (BRR) technique, Fay's Method, should be used. Sample SAS code is listed below:

```
proc surveyfreq data=work.mcbs VARMETHOD = brr (fay=.30);
tables sex*d_strat/ col notot;
weight cs1yrwgt;
repweights cs1yr001-cs1yr100;
run;
```

To compute the standard errors with MCBS cross sectional weights and design variables, Taylor series linearization should be used.[1] Sample SAS code is listed below:

proc surveyfreq data=work.mcbs;

tables sex*d_strat/ col notot;

strata sudstrat;

cluster sudunit;

weight cs1yrwgt;

run;

- **How can I conduct subgroup analyses and maintain the appropriate variance estimation?**

  Variance estimation can be impacted by selecting individuals prior to analysis. 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. 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 VARMETHOD = brr (fay=.30);

  table sex * diabetes * flag / col notot;

  weight cs1yrwgt;

  repweight cs1yr001-cs1yr100;

  run;

  ….where the FLAG variable is set to 1 if the beneficiary is over 65 and in Medicare Advantage, 0 otherwise (for example). Basically, you want to keep the full population in your analysis dataset for strictly accurate variance estimation. Point estimates will not be affected.

Segment Guidance – Survey File

Demographics (DEMO)

- **How was INT_TYPE created?**

  INT_TYPE is defined as:

  - C = respondent only resided in the community and only completed Community-administered survey instruments in each round for 2015
  - F = respondent only resided in a facility and only completed Facility-administered survey instruments in each round for 2015

[1] Given the rotating panel design of the MCBS, performing pooled cross-sectional analysis using Taylor-Series Linearization method of variance estimation will require additional adjustments to account for non-independence of beneficiaries across years in a multi-year dataset.

- B = respondents completed instruments in both settings across the rounds for 2015

INT_TYPE was created following the rules below:

- For all cases that completed an interview in both the 2015 Winter/Summer and Fall rounds, INT_TYPE = B if the setting in which the survey instruments was completed differs between the Winter/Summer and Fall rounds. All other beneficiaries who completed survey instruments across the same settings in both the Winter/Summer and Fall rounds were assigned either C or F accordingly.
- For any beneficiaries that did not complete interviews in both rounds, then the INT_TYPE reflected the setting the beneficiary resided in during the Fall 2015 round.
- For any beneficiaries still missing INT_TYPE, then INT_TYPE was assigned as the instrument completed in the 2015 Winter/Summer round.

Missing INT_TYPEs - There are currently 42 beneficiaries with "complete" dispositions which cannot have their INT_TYPE/residence location calculated for them in 2015. These are individuals who died in early 2015, and did not have any complete questionnaire data for 2015. These individuals have ever enrolled weights, but do not have complete interviews. Note that in each data year, some differences by segment will exist (i.e., data may reflect a prior or future calendar year due to the specific questionnaire and reference period used to collect the information).

Facility Assessments (FACASMNT)

- **What is the difference between the MDS and FACASMNT data?** See exhibit below for key differences between the segment sources, population, reference period and unit of observation.

Exhibit 1: Differences between FACASMNT and MDS3 Data

Data Type	Facility Assessment (FACASMNT)	Minimum Data Set (MDS3)
Source	Survey-reported (staff may pull from their electronic health records or systems to answer the survey questions)	Administrative (like claims files)
Population	Represents ALL facility residents, not just those in nursing homes	Represents all **residents of nursing homes certified to participate in Medicare or Medicaid** ONLY
Reference period	Throughout the year	Could be multiple assessments during the year, time periods may differ based on what happened to each individual
Unit of observation	One per beneficiary	One per beneficiary per assessment

Food Insecurity (FOODINS)

- Please see the Weights section below for information on using weights with data from Topical Questionnaires.
- There are two to four Facility-dwelling beneficiaries (INT_TYPE = F) that appear on the 2015 segments that include 2016 non-response adjustments (i.e., INCASSET, FOODINS, MCREPLNQ, RXPARTD, PNTACT) due to the fact that the data was collected in 2016, but INT_TYPE is created using 2015 residence only.

Health Insurance Summary (HISUMRY)

- **Medicare/Medicaid dual eligibility indicator (H_OPMDCD)** – In 2015, CMS modified the data types of several variables. One example is the 2013 variable OP_MDCD compared to the 2015 variable H_OPMDCD. The variable values "1","2","3","4" had been stored as character values. As part of the modifications applied in 2015, these values are stored as numeric values. The code definitions are equivalent. Users that want to merge the HISUMRY segment across Access to Care (2013 and prior) and Survey File (2015) must change the data type of one variable and rename it.
- **Does Group Health Participation (H_GHPSW) refer to MA?** Some of the beneficiaries in the MCBS sample belong to Medicare managed care plans. CMS derived variables that describe this Medicare managed care membership (H_GHPSW and MAFLAG01 - MAFLAG12). The variable (H_GHPSW) can be used when only an indication that the enrollee was a member of a Medicare managed care plan at some time during 2015 is needed for analysis. The monthly variables (MAFLAG01- MAFLAG12) can be used for analyzing membership at specific points in time. The variables will indicate either "FFS" (Fee for Service) or "MA" (Medicare Advantage).

  The H_GHPSW variable is derived from the Health Maintenance Organization (HMO) Coverage Months variable in the administrative data. This variable indicates participation in a group health organization, also known as HMO, managed-care participation, or Medicare Advantage.

- **What happens to missing values in TOT_PREM?** Prior to 2011 we would set all of the missing values for TOT_PREM to 0. In 2011, we left unreported (missing) as missing (.). Setting the missing values to 0 would provide comparable numbers of 0 to previous year trends.

  How do I identify which beneficiaries have private managed care plans and employer specialty plans? The Health Insurance Timeline (HITLINE) and Health Insurance Summary (HISUMRY) segments include flags/codes for Private Managed Care Plan and Employer-sponsored (ESI) Specialty Plans. In HISUMRY, S_TYPPL1-S_TYPPL5 = 14 indicates whether the beneficiary had an ESI Specialty Plan. S_TYPPL1-S_TYPPL5 = 13 indicates whether the beneficiary had a self-purchased private HMO/managed care plan.

Health Insurance Timeline (HITLINE)

- The 2015 HITLINE segment is provided in a long format versus a wide format. If data users are conducting analyses across multiple years, then they will need to transpose the data in order to merge.
- This segment contains the types of insurances, the coverage eligibility timeline, and the source information for the coverage. HITLINE contains one record for each plan type a beneficiary has, and then provides information on the monthly eligibility/enrollment and the source of that data. The segment also includes information on Part D coverage. PLANTYPE = 4 is Medicare Part D/Part D/MAPD.
- **What happened to variables H_PDTP01-H_PDTP12?** The variables H_PDTP01 – H_PDTP12 from the Cost & Use RIC A were unreliable, and beginning with CAU 2011 these monthly Part D plan type variables were removed. In lieu of these variables, there is now a Part D plan flag indicator (PLANTYPE = 4) and monthly coverage and eligibility indicators on the HITLINE segment.
- **How do I identify which beneficiaries have private managed care plans and employer specialty plans?** The Health Insurance Timeline (HITLINE) and Health Insurance Summary (HISUMRY) segments include flags/codes for Private Managed Care Plan and Employer Specialty Insurance (ESI) Plans. In HITLINE, PLANTYPE = 14 indicates whether the beneficiary had an ESI Plan.
  PLANTYPE = 13 indicates whether the beneficiary had a self-purchased private HMO/managed care.

Income and Assets (INCASSET)

- Please see the Weights section below for information on using weights with data from Topical Questionnaires.
- There are two to four Facility-dwelling beneficiaries (INT_TYPE = F) that appear on the 2015 segments that include 2016 non-response adjustments (i.e., INCASSET, FOODINS, MCREPLNQ, RXPARTD, PNTACT) due to the fact that the data was collected in 2016, but INT_TYPE is created using 2015 data only.

Interview Characteristics (INTERV)

- **What does the INTVDISP variable represent**? INTVDISP is the respondent's status as of the last day of the calendar year. There are two dispositions (40 = complete and 50 = complete-deceased). This variable was called STATUS in the 2013 CAU LDS release and is now found on the Survey File LDS' INTERV segment.

Medicare Plan Beneficiary Knowledge (MCREPLNQ)

- Please see the Weights section below for information on using weights with data from Topical Questionnaires.
- There are two to four Facility-dwelling beneficiaries (INT_TYPE = F) that appear on the 2015 segments that include 2016 non-response adjustments (i.e., INCASSET, FOODINS, MCREPLNQ, RXPARTD, PNTACT) due to the fact that the data was collected in 2016, but INT_TYPE is created using 2015 data only.

Minimum Data Set (MDS3)

- **What is the difference between the MDS and FACASMNT data?** See Facility Assessments (FACASMNT) section above.
- **Why are there community-dwelling beneficiaries (INT_TYPE = 'C') represented on the MDS3 segment?**
  - There are Community-dwelling beneficiaries (DEMO segment INT_TYPE = C) that appear in the MDS segment. CMS includes MDS data for all MCBS beneficiaries regardless of the INT_TYPE, which is determined by the type of survey instrument completed.
  - The MDS is administered by any qualified Medicare facility, including Skilled Nursing Facilities (SNF). A beneficiary may have had one or more SNF stays during the course of the year, but still resided in the community and completed their interviews there.

Part D Drug Plan Experience (RXPARTD)

- Please see the Weights section below for information on using weights with data from Topical Questionnaires.
- There are two to four Facility-dwelling beneficiaries (INT_TYPE = F) that appear on the 2015 segments that include 2016 non-response adjustments (i.e., INCASSET, FOODINS, MCREPLNQ, RXPARTD, PNTACT) due to the fact that the data was collected in 2016, but INT_TYPE is created using 2015 data only.

Patient Activation (PTNTACT)

- Please see the Weights section below for information on using weights with data from Topical Questionnaires.
- There are two to four Facility-dwelling beneficiaries (INT_TYPE = F) that appear on the 2015 segments that include 2016 non-response adjustments (i.e., INCASSET, FOODINS, MCREPLNQ, RXPARTD, PNTACT) due to the fact that the data was collected in 2016, but INT_TYPE is created using 2015 data only.

Residence Timeline (RESTMLN)

- For data year 2015 only, RESTMLN is included with the Cost Supplement File since it was constructed using the 2015 Cost Supplement population, and should be analyzed using the 2015 Cost Supplement weights. For data year 2016 and beyond, the segment will be included with the Survey File.

Weights

- **How should I use weights with data from Topical Questionnaires?** To generate estimates using the data from one of the five Topical Questionnaire segments (FOODINS, INCASSET, MCREPLNQ, PNTACT, RXPARTD) on their own or merged with another Survey File segment that does not contain special non-response adjustment weights, the analyst must always use the special non-response adjustment general and replicate weights included in the Topical segment instead of using the general and replicate weights that appear in the separate weight segments (CENWGTS, EVRWGTS).

  There are no weights that support joint analysis between two topical modules. Each segment with data from Topical Questionnaires has a different set of beneficiaries included. A user could merge data from one Topical segment onto another and then use one of the Topical segment's weights as the Baseline population, but the availability of data will not line up and there will be gaps. For some combinations of the different questionnaire sections, the amount of missing data may be small enough that users could still conduct analyses.

Segment Guidance – Cost Supplement File

- **Can I use the Survey File longitudinal weights with the 2015 Cost Supplement data?** The 2015 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.
- **What does survey-reported DME information include?** While we do not use the term "Durable Medical Equipment" or "DME" in the Community questionnaire, there are questions in the Other Medical Expenses (OMQ) section of the questionnaire that cover most common DMEs, such as blood sugar monitors, blood sugar testing strips, canes, crutches, hospital beds, wheelchairs, nebulizers, sleep apnea and continuing positive airway pressure (CPAP) devices, oxygen equipment and accessories, and walkers. The interviewers are not expected to determine which equipment count as a DME or not—they are expected to record the correct equipment type at the appropriate question in the OMQ section.

  For most items reported in the OMQ, the questionnaire collects item type, dates of purchase, and associated costs and payments. The questionnaire is designed to capture any equipment purchases, regardless of whether the equipment was purchased from medical providers or directly from the retail

outlet and regardless of payment source. Information about payment is gathered during the charge series. The questionnaire does not directly collect information about place of purchase.

- **How is Part B drug info collected? Where is this located in the research files?** A few Part B drugs are collected as survey reports in the PME. However, the data we add from claims is only from Part D, so we would not be attempting to match any survey reported drugs administered by a physician to the actual Part B administrative claims data.
- **How can I determine whether a drug is brand name or generic?** A very non-scientific way to determine brand vs. generic is if the FDB_BN field is different than the FDB_GNN field, it is potentially a brand name drug (or at least has a trademarked name).

  Data users can also use the PDE NDC and use an external drug information database (like FDB) and use some other fields to determine brand vs. generic status of the drug.

  If data users do not have a drug database, users could use the Food and Drug Administration's NDC SPL DataElements File. CMS uses this for the Manufacturer drug discount program to determine what products are not eligible for the Manufacturer discount. Based on the Marketing category, drugs can be classified as follows:

  Brand: NDA. NDA authorized generic, BLA
  Generic: ANDA