

USER GUIDE
Medicare Current Beneficiary Survey
2013 Access to Care, Public Use File

1. Introduction

The Medicare Current Beneficiary Survey (MCBS) 2013 Access to Care (ATC) Public Use File (PUF) was developed in response to needs expressed as part of the Centers for Medicare & Medicaid Services (CMS) sponsored Alliance to Modernize Healthcare (CAMH). The CAMH project, while supporting the value proposition of the MCBS, highlighted several barriers in using MCBS data for research and program evaluation:

- The costs of the files;
- The Data Use Agreement process;
- The complexity of the data structures and,
- The fact that researchers cannot determine until after they have purchased the files what research they would like to conduct, and if the information needed is available.

The MCBS was only available to researchers as a Limited Data Set (LDS). An important suggestion outlined in the report to address these barriers, was the development of a streamlined de-identified public-use file, which would improve data access, lead to more and different types of researchers using the MCBS and allow preliminary research to be conducted.

As part of the Administration's goals of better care, smarter spending, and healthier people, CMS is committed to increasing access to its data. Over the past several years, CMS has made it a priority to make more data available, including releasing to the public an unprecedented amount of information on services and procedures provided to Medicare beneficiaries. The development of the MCBS PUF was supported by the CMS Office of Minority Health as part of their activities to increase the availability of data to monitor the impact of CMS programs in improving minority health and eliminating health disparities. As part of these larger goals, CMS is releasing the MCBS 2013 ATC PUF.

The MCBS 2013 ATC PUF is not intended to replace the more detailed LDS files, rather it will provide a publically available alternative for those researchers interested in the health, health care use, access to and satisfaction with Medicare of beneficiaries, while providing the very highest degree of protection to the Medicare beneficiaries' protected health information. The main benefits provided to potential researchers of the MCBS 2013 ATC PUF are:

1. Increased data access through a free file download that is consistent with other HHS public-use survey files;
2. Increase policy-relevant analyses, by attracting new researchers and policy-makers, for whom the cost and time associated with accessing the MCBS LDS can pose significant deterrents to use;
3. Promote research to improve the Medicare program; and
4. Provide publically available data for key indicators of the Medicare population for benchmarking comparisons with other data sources.

2. Background of MCBS

The MCBS, sponsored by the CMS in partnership with the Center for Medicare and Medicaid Innovation (CMMI) and the Office of Enterprise and Data Analytics (OEDA), is a continuous, in-person, longitudinal survey of a representative national sample of the Medicare population, covering the population of beneficiaries in the US, District of Columbia, and Puerto Rico. It has been carried out continuously for more than 20 years, encompassing more than one million interviews. The MCBS is designed to aid CMS in administering, monitoring, and evaluating Medicare programs, is the leading source of information on Medicare and its impact on beneficiaries, provides important information on Medicare beneficiaries that is NOT available in CMS administrative data and plays an essential role in monitoring and evaluating key provisions of the Affordable Care Act.

3. Technical and programming information

General Information: The MCBS 2013 ATC PUF includes 13,924 community interview survey respondents. All records begin with a PUF_ID, a unique number for each beneficiary in the public use file. This PUF_ID serves to identify records in the PUF only and cannot be used for linking to other sources of data.

All variables on the PUF are numeric (except for survey year) or integer. Formats and values for each variable are available in the PUF codebook that is available with the data download.

Variable groups are labeled with prefixes to help users identify these groups by topic area and the associated questionnaire sections and original file locations (the 2013 ATC individual files called RICs). Table 1 includes information about these variable prefixes.

Table 1: MCBS 2013 ATC PUF Variable Prefix

Prefix	Descriptions	LDS RIC Variable source
IV_	Age, sex, and race groups	RIC1
D_	Other demographic variables	RIC1
HLT_	Health Status/Functioning (Community)	RIC2
PRV_	Health Status/Functioning (Prevention)	RIC2P
ACC_	Access to, use, and satisfaction with health care	RIC3
INS_	Insurance status, coverage, and type	RIC4
HOU_	Housing characteristics	RIC5
INT_	Interview characteristics	RIC8
MA_	HMO Supplement	RICH
ADM_	Administrative Data	RICA
CS	Survey weights	RICX

Data File Information: All information about the MCBS 2013 ATC PUF dataset can be found in the codebook. The codebook includes SAS variable names, labels, a universe statement which indicates which respondents were eligible for the questions and the questionnaire text that was asked during the survey. For certain variables some of the questionnaire text does not match the categories provided in the public release file because certain categories were recoded due to disclosure concerns (e.g. “no usable vision” for variable HLT_ECTROUB and “deaf” for HLT_HCTROUB). Other variables were created based on combining two questions and their questionnaire text indicates a recoded variable (e.g. HLT_ALZDEM). Certain variables combine survey reported data and administrative data (e.g. INS_D_MCARE), their questionnaire text indicates “Recode of administrative data and survey-reported data.”

For each variable, the formats and format values are included in the codebook. Values of -7 indicate “refuse” and -8 indicate “don’t know.” All values of “inapplicable” have been combined with missing values. Both weighted and unweighted frequencies of most variables included in MCBS 2013 ATC PUF are provided in the accompanying codebook file.

The MCBS 2013 ATC PUF dataset is saved as a SAS export file. Directions and sample SAS code are given below to help users read the dataset into SAS.

Assume you have downloaded the MCBS 2013 ATC PUF export (ATCPUF13.xpt) file in the folder "C:\MCBS\DOWNLOAD". You can then use the following SAS code to import it into SAS.

```
LIBNAME PUFLIB 'C:\MCBS\SASDATA';  
FILENAME F "C:\MCBS\DOWNLOAD\ATCPUF13.XPT";  
  
PROC CIMPORT LIBRARY=PUFLIB INFILE=F;  
RUN;
```

Similarly, a text file with SAS programming code to create formats is provided for users.

Comparison to LDS: The MCBS 2013 ATC PUF differs from 2013 ATC LDS files because it has been evaluated for disclosure risk and additional steps were taken to protect respondent confidentiality. 950 observations and 1,284 variables were dropped. Additionally, variables were recoded or recreated based on variables included in the LDS.

The PUF includes only those beneficiaries interviewed in the community and, thus, excludes all beneficiaries who were in a health care facility part or all of the year (n=950). Variables that were only created for facility residents were also not included. Facilities are defined as nursing homes, retirement homes, domiciliary or personal care facility, distinct long term units in a hospital complex, mental health facility and centers.

Additionally, the PUF contains no Medicare claims data in individual form. For users interested in this type of data, the MCBS 2013 ATC LDS is available from CMS through a data use agreement. More information on the LDS process can be found here:

<https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS/>.

A summary of the differences between the two data products appears in Table 2 below:

Table 2: Comparison Between the MCBS 2013 ATC PUF and MCBS 2013 ATC LDS

MCBS PUF	MCBS LDS
Population – community dwelling	Population – community and facility
Number of variables - 459	Number of variables – 1,742
ID – new PUFID Randomly generated, can’t be linked back to BASEID	ID – BASEID Randomly generated, can’t be linked back to HIC number
Date fields - NO	Date fields - YES
Geographic identifiers – NO	Geographic identifiers - YES
Cost/payment data – NO	Cost/payment data – YES
Demographic data – YES All variables are categorical	Demographic data – YES Continuous age variables available
Insurance coverage – YES Summarized to annual level	Insurance coverage – YES Monthly level
Identifiable plan-related information for MA or Part D – NO	Identifiable plan-related information for MA or Part D - YES

4. Survey sample information

Background on design and response rates: Respondents for the MCBS were sampled from the Medicare enrollment file to be representative of the Medicare population as a whole and by the following age groups: under 45, 45 to 64, 65 to 69, 70 to 74, 75 to 79, 80 to 84, and 85 and over. Because of interest in their special health care needs, the oldest old (85 and over) and the disabled (64 and under) were oversampled to permit more detailed analysis of these subpopulations. Beginning in 2013, an additional oversample of beneficiaries in Accountable Care Organizations (ACOs) was implemented. The MCBS sample was designed to yield about 16,000 completed cases annually in the Access to Care data file including those residing in facilities. The MCBS 2013 ATC PUF excludes facility residents.

The beneficiaries included in the 2013 Access to Care File consist of a random cross-section of all beneficiaries who were continuously enrolled in one or both parts of the Medicare program from January 1, 2013 up to and including their interview during the 2013 fall round (September - December). These beneficiaries include those in four separate MCBS longitudinal panels identified by the year in which the panel was selected (i.e., the 2010, 2011, 2012 and 2013 panels) and were drawn through the use of a fairly complex selection algorithm.

The 2013 MCBS data are a point in time reflection of Medicare beneficiaries residing in the United States or Puerto Rico who were enrolled in one or both parts of the program. This “always enrolled” population represents individuals continuously enrolled from January 1, 2013 up to and including their fall interview. Excluded are the following categories of Medicare enrollees:

- 1) Residents of foreign countries and U.S. possessions and territories other than Puerto Rico;
- 2) Persons who became enrolled after January 1, 2013; and
- 3) Persons who dis-enrolled or died prior to their Fall interview;

The MCBS sample is designed to provide nearly self-weighting (i.e., equiprobability) samples of beneficiaries within each age stratum. Within the selected PSUs and ZIP clusters, a systematic sampling scheme with random starts is employed. A brief summary of the 2010 – 2013 panels is given below.

In 2010 a supplemental sample of 7,260 beneficiaries was added to the sample for Round 58. The 2010 supplemental MCBS sample included newly enrolled beneficiaries (i.e., beneficiaries who were enrolled during the period February 2009 through January 2010) as well as previously enrolled beneficiaries (i.e., beneficiaries who were enrolled on or before January 2009). The 2010 Access to Care questions were administered in September through December 2010 as part of the Round 58 interview for the continuing sample, and as part of the initial interview for the 2010 supplemental sample.

In 2011 a supplemental sample of 7,365 beneficiaries was added to the sample for Round 61. The 2011 supplemental MCBS sample included newly enrolled beneficiaries (i.e., beneficiaries who were enrolled during the period February 2010 through January 2011) as well as previously enrolled beneficiaries (i.e., beneficiaries who were enrolled on or before January 2010). The 2011 Access to Care questions were administered in September through December 2011 as part of the Round 61 interview for the continuing sample, and as part of the initial interview for the 2011 supplemental sample.

In 2012 a supplemental sample of 7,400 beneficiaries was added to the sample for Round 64. The 2012 supplemental MCBS sample included newly enrolled beneficiaries (i.e., beneficiaries who were enrolled during the period February 2011 through January 2012) as well as previously enrolled beneficiaries (i.e., beneficiaries who were enrolled on or before January 2011). The 2012 Access to Care questions were administered in September through December 2012 as part of the Round 64 interview for the continuing sample, and as part of the initial interview for the 2012 supplemental sample.

Finally, in 2013 a supplemental sample of 7,400 beneficiaries was added to the sample for Round 67. The 2013 supplemental MCBS sample included newly enrolled beneficiaries (i.e., beneficiaries who were enrolled during the period February 2012 through January 2013) as well as previously enrolled beneficiaries (i.e., beneficiaries who were enrolled on or before January 2012). The 2013 Access to Care questions were administered in September through December 2013 as part of the Round 67 interview for the continuing sample, and as part of the initial interview for the 2013 supplemental sample.

The cross-sectional sample in any given year or round is then composed of members of all of the eligible panels. The figure below illustrates the panel composition for the 2013 MCBS ATC PUF.

Figure 1: MCBS Panels in the MCBS 2013 ATC PUF

	2010			2011			2012			2013			2014		
	56 Jan- Apr	57 May- Aug	58 Sep- Dec	59 Jan- Apr	60 May- Aug	61 Sep- Dec	62 Jan- Apr	63 May- Aug	64 Sep- Dec	65 Jan- Apr	66 May- Aug	67 Sep- Dec	68 Jan- Apr	69 May- Aug	70 Sep- Dec
Panel 2010															
Panel 2011															
Panel 2012															
Panel 2013															
												ATC 2013			

As with most data collection activities, some existing and new panel members will either be impossible to locate or will refuse to participate in an interview. The calculation of the study-wide response rates generally follow the guidelines specified in AAPOR¹ and OMB². For the cross-sectional sample represented by the MCBS 2013 ATC PUF, the calculated overall response rate was 55.6%. This rate includes non-response for persons in facilities, as the response rates are not calculated separated by questionnaire component. Therefore, this may not reflect exactly the response rate for the sample represented in MCBS 2013 ATC PUF which excludes beneficiaries in a facility.

Weights and variance estimation: The sample design of MCBS includes stratification, clustering, multiple stages of selection, and disproportionate sampling. Furthermore, the MCBS sampling weights reflect adjustments for survey nonresponse. These survey design and estimation complexities require special consideration when analyzing MCBS data (i.e., it is not appropriate to assume simple random sampling).

To obtain accurate estimates from MCBS data, for either descriptive statistics or more sophisticated analyses based on multivariate models, the survey design complexities need to be taken into account by applying MCBS weights to produce estimates and using an appropriate technique to derive standard errors associated with the weighted estimates.

The full sample weight is CS1YRWGT. CS1YRWGT is the “always enrolled” cross-sectional weight and applies to both the continuing sample and to the supplemental sample that was fielded for the first time in Round 67. This weight is intended for use in cross-sectional statistics involving the total (combined) Round 67 sample. Each weight is greater than zero for all beneficiaries on the file. CS1YRWGT should be used to make estimates of the levels of access to care for the Medicare population alive in the fall of 2013 (i.e., the always-enrolled population).

To permit the calculation of sampling errors, a series of replicate weights were computed. Unless the complex nature of the MCBS is taken into account, estimates of the variance of a survey statistic may be biased downward. The replicate weights included in the MCBS 2013 ATC PUF can be used to calculate standard errors of the sample-based estimates as described below. The replicate cross-sectional weights are labeled CS1YR001 through CS1YR100 corresponding to the always-enrolled weight CS1YRWGT.

¹ https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions2015_8theditionwithchanges_April2015_logo.pdf

² http://www.whitehouse.gov/sites/default/files/omb/inforeg/statpolicy/standards_stat_surveys.pdf

Most commercial software packages today include techniques to accommodate the complex design, through replicate weight approaches. Among these are STATA®, SUDAAN®, R®, and the complex survey procedures in SAS®. When using the replicate weight approach to variance estimation, we recommend the variance estimation method of balanced repeated replication using Fay's adjustment of 0.3. Sample code in SAS and STATA for estimating statistics follow below.

5. Sample Code

SAS ANALYSIS STATEMENTS

Cross-tabulations

```
proc surveyfreq data=<Analytic dataset> VARMETHOD = brr (fay=.30);
```

```
    table <Var name> / row chisq lrchisq ;
```

```
    weight cs1yrwgt;
```

```
    repweight cs1yr001 - cs1yr100;
```

```
run;
```

Mean / Sum

```
proc surveymeans data=<Analytic dataset> plots=NONE sum VARMETHOD = brr (fay=.30);
```

```
    var <Var name>;
```

```
    weight cs1yrwgt;
```

```
    repweight cs1yr001 - cs1yr100;
```

```
run;
```

Subgroup Analysis

```
proc surveymeans data=<Analytic dataset> plots=NONE sum VARMETHOD = brr (fay=.30);
```

```
    var <Var name>;
```

```

weight cs1yrwgt;

repweight cs1yr001 - cs1yr100;

domain <Subgroup variable>;

run;

proc surveyfreq data=<Analytic dataset> VARMETHOD = brr (fay=.30);

    table <Var name> * <Subgroup variable> / row chisq lrchisq ;

weight cs1yrwgt;

repweight cs1yr001 - cs1yr100;

run;

```

STATA ANALYSIS STATEMENTS

Declare dataset as survey sample with replicate weights

```
svyset _n [pweight=cs1yrwgt], brrweight(cs1yr001-cs1yr100) fay(.3) vce(brr) singleunit(missing)
```

For categorical variables, use:

```
svy brr, fay(.3) : tabulate <Var name> <Var name>
```

For continuous variables, use:

```
svy brr, fay(.3) : mean <Var name>, over(<Var name>)
```

For subgroup analysis use:

```
svy brr, subpop(if <Subgroup>) fay(.3) : mean <Var name>, over(<Var name>)
```