# MCBS
## Medicare Current Beneficiary Survey

# 2021 | DATA USER'S GUIDE: SURVEY FILE PUBLIC USE FILE

Centers for Medicare & Medicaid Services (CMS)
Office of Enterprise Data and Analytics (OEDA)

# Version Control Log

| Date | Version | Revisions |
|------|---------|-----------|
| 10/13/2023 | 1.0 | Initial version published. |

# OVERVIEW OF MCBS DOCUMENTATION

The Centers for Medicare & Medicaid Services (CMS) releases a comprehensive suite of documentation products to support researchers in using the Medicare Current Beneficiary Survey (MCBS). This section provides a concise summary of each documentation product.

- **Data User's Guides**: A Data User's Guide is produced for each MCBS Limited Data Set (LDS) and Microdata Public Use File (PUF) data release. There are three broad categories of Data User's Guides.

  - *Survey File Data User's Guide*: Updated annually for each new data year, the *Survey File Data User's Guide* supports researchers in understanding and analyzing Survey File LDS data. This Data User's Guide contains detailed information about the Survey File LDS, including changes between years, important data user considerations, and sample code, as well as basic background information on the MCBS, including sampling, questionnaires, data collection, and data processing. Along with the *New User Tutorial* (see below), this Data User's Guide is the recommended starting point for researchers, particularly for those new to studying MCBS data.

  - *Cost Supplement File Data User's Guide*: Updated annually for each new data year, the *Cost Supplement File Data User's Guide* functions as a supplement to the corresponding *Survey File Data User's Guide* and supports researchers in understanding and analyzing Cost Supplement File LDS data. This Data User's Guide focuses on providing detailed information about the Cost Supplement File LDS, including changes between years, important data user considerations, and sample code.

  - *Public Use File Data User's Guides* (this document): A Data User's Guide is also produced for each MCBS Microdata PUF release, including the annual Survey File PUF, the annual Cost Supplement File PUF, and the three COVID-19 Supplement PUFs. These Data User's Guide provide detailed, focused information to support researchers in understanding and analyzing PUF data.

- **Methodology Report**: Updated annually for each new data year, the *Methodology Report* provides detailed background information on the methods used to conduct the MCBS and process MCBS data. This includes information on sampling methodology, questionnaire development and programming, interviewer recruitment and training, data collection procedures, data processing and editing, including weighting and imputation, and response rates.

- **Data User Tutorials**:

  - *New User Tutorial*: Aimed at new data users who are unfamiliar with the MCBS, the *New User Tutorial* provides an overview of MCBS history, policy relevance, survey design, data products, and best practices for analysis. Along with the *Survey File Data User's Guide* (see above), the *New User Tutorial* is the recommended starting point for researchers.

  - *Advanced Topic-Based Tutorials*: In addition to the *New User Tutorial*, CMS has released a series of tutorials on more advanced topics, with the goal of supporting researchers in better understanding how to analyze and interpret MCBS data by providing detailed analytic guidance and examples. Topics of these tutorials include the differences between MCBS Community and Facility data, weighting and variance estimation, using data from the MCBS COVID-19 Supplements, conducting longitudinal analysis, and conducting pooled cross-sectional analysis with MCBS data.

## TABLE OF CONTENTS

## LIST OF EXHIBITS

# ACRONYM LIST

| | |
|---|---|
| AAPOR | American Association for Public Opinion Research |
| ACCESSCR | LDS Survey File Access to Care segment |
| ACCSSMED | LDS Survey File Access to Care, Medical Appointments segment |
| ADMNUTLS | LDS Survey File Administrative Utilization Summary segment |
| ASSIST | LDS Survey File Assistance segment |
| CAPI | Computer-Assisted Personal Interviewing |
| CHRNCOND | LDS Survey File Chronic Conditions segment |
| CMS | Centers for Medicare & Medicaid Services |
| COMMDOSE | LDS Survey File Community Questionnaire COVID-19 Vaccine Dosage segment |
| CSV | Comma-separated values file |
| DEMO | LDS Survey File Demographics segment |
| DUA | Data Use Agreement |
| EVRWGTS | LDS Survey File Ever Enrolled Population Weight segment |
| FALLS | LDS Survey File Falls segment |
| FOODINS | LDS Survey File Food Insecurity segment |
| GENHLTH | LDS Survey File General Health segment |
| HHCHAR | LDS Survey File Household Characteristics segment |
| HHS | U.S. Department of Health and Human Services |
| HIC | Health Insurance Claim |
| HISUMRY | LDS Survey File Health Insurance Summary segment |
| HITLINE | LDS Survey File Health Insurance Timeline segment |
| INTERV | LDS Survey File Interview Characteristics segment |
| IRB | Institutional Review Board |
| LDS | Limited Data Set(s) |
| MA | Medicare Advantage |
| MAPLANQX | LDS Survey File Medicare Advantage Plan Questions segment |
| MCREPLNQ | LDS Survey File Medicare Plan Beneficiary Knowledge segment |
| MCBS | Medicare Current Beneficiary Survey |
| NAGIDIS | LDS Survey File Nagi Disability segment |
| NICOALCO | LDS Survey File Nicotine and Alcohol segment |
| NORC | NORC at the University of Chicago |
| OMB | Office of Management and Budget |
| PHI | Protected Health Information |
| PII | Personally Identifiable Information |
| PREVCARE | LDS Survey File Preventive Care segment |
| PSU | Primary Sampling Units |
| PUF | Public Use File |
| RXMED | LDS Survey File RX Medications segment |
| SAS | Statistical Analysis System |
| SATWCARE | LDS Survey File Satisfaction with Care segment |
| SSU | Secondary Sampling Units |
| TELEMED | LDS Survey File Telemedicine segment |
| USCARE | LDS Survey File Usual Source of Care segment |
| USU | Ultimate Sampling Unit |
| VISHEAR | LDS Survey File Vision and Hearing segment |

# 1. INTRODUCTION

Over the past several years, the Centers for Medicare and Medicaid Services (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. CMS provides users with multiple ways to access Medicare Current Beneficiary Survey (MCBS) data, and a wide array of documentation is publicly available on the CMS MCBS website. MCBS data are made available via two annual Limited Data Set (LDS) releases, and two annual Microdata Public Use File (PUF) releases, an MCBS Survey File PUF based on the Survey File LDS and an MCBS Cost Supplement File PUF based on the Cost Supplement File LDS.[1] In addition, in response to the emergence of "coronavirus disease 2019" ("COVID-19") in the United States in 2020, the MCBS has also released three out-of-cycle topic-specific Microdata PUFs (MCBS COVID-19 Summer 2020 PUF, MCBS COVID-19 Fall 2020 PUF, and MCBS COVID-19 Winter 2021 PUF).

The content of the MCBS Survey File PUF is governed by its central focus of serving as a unique source of information on beneficiaries' health and well-being that cannot be obtained through CMS administrative sources alone. The file includes data related to Medicare beneficiaries' access to care, health status, other information regarding beneficiaries' knowledge of, attitudes toward, and satisfaction with their health care, as well as demographic data and information on all types of health insurance coverage. Disclosure protections have been applied to the file, including de-identification and other methods; as a result, the MCBS Survey File PUF does not require a Data Use Agreement (DUA). In contrast, the MCBS LDS releases contain beneficiary-level protected health information (PHI) and therefore require a DUA. The MCBS Survey File PUF is not intended to replace the more detailed LDS files; rather, it makes available a general-use publicly-available alternative that provides the highest degree of protection to the Medicare beneficiaries' PHI.

The main benefits of the MCBS Survey File PUF are:

1. Increased data access for researchers of the MCBS through a free file download that is consistent with other U.S. Department of Health and Human Services (HHS) public-use survey files;

2. Increased 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.

This user guide contains information about the 2021 MCBS Survey File PUF. It contains detailed information about the MCBS and specific background information to help data users understand and analyze the PUF. This guide is updated each time a new set of PUF data are released.[2]

Readers interested in understanding or analyzing the 2021 MCBS data should also familiarize themselves with the content of the *2021 Data User's Guide: Survey File* and the *2021 MCBS Methodology Report* documents in order to obtain an overview of the survey, questionnaires, sample design, and other topics relevant to the MCBS. Data users can access these documents along with other data documentation at: https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS/Codebooks. Data users interested in a collection of charts and tables presenting estimates from the LDS releases can access the *MCBS Chartbook* at: https://www.cms.gov/Research-Statistics-Data-andid a d-Systems/Research/MCBS/Data-Tables.

---

[1] The MCBS Cost Supplement File PUF is based on the MCBS Cost Supplement File LDS; however, this PUF also includes select demographic and health factor variables from the MCBS Survey File LDS.

[2] This communication was printed, published, or produced and disseminated at U.S. taxpayer expense.

# What's New in 2021?

Below, data users will note highlights and updates to the MCBS Survey File PUF for the 2021 data year. Detailed information about changes to the 2021 MCBS, sampling, questionnaires, documentation, and data processing is available in the *2021 MCBS Data User's Guide: Survey File.*

## 1.1    What's New in the 2021 MCBS Survey File PUF?

The 2021 MCBS Survey File PUF features the following updates:

Two arthritis variables, which had appeared in previous versions of the MCBS PUF prior to 2019, were re-added to the 2021 Survey File PUF Fall segment. In addition, two variables were removed from the 2021 Survey File PUF Fall segment about time since last colonoscopy and time since last oral exam.

Three variables were added to the 2021 Survey File PUF Winter segment about owning or using a desktop or laptop computer, a smartphone, and a tablet or portable computer, while one variable was removed about having a personal computer. Six variables about telehealth were also added to the 2021 Survey File PUF Winter segment. In addition, 15 variables about seeing a doctor were dropped from the 2021 Survey File PUF Winter segment.

One variable about COVID-19 vaccination was added to the 2021 Survey File PUF Summer segment.

Exhibit 1.1.1 presents a comparison between the number of variables in the MCBS Survey File PUF for data years 2013, 2015, 2016, 2017, 2018, 2019, 2020, and 2021 including weights for the seasonal segments.

---

**Exhibit 1.1.1:**    Number of Variables in the 2013, 2015, 2016, 2017, 2018, 2019, 2020, and 2021 MCBS Survey File PUF

---
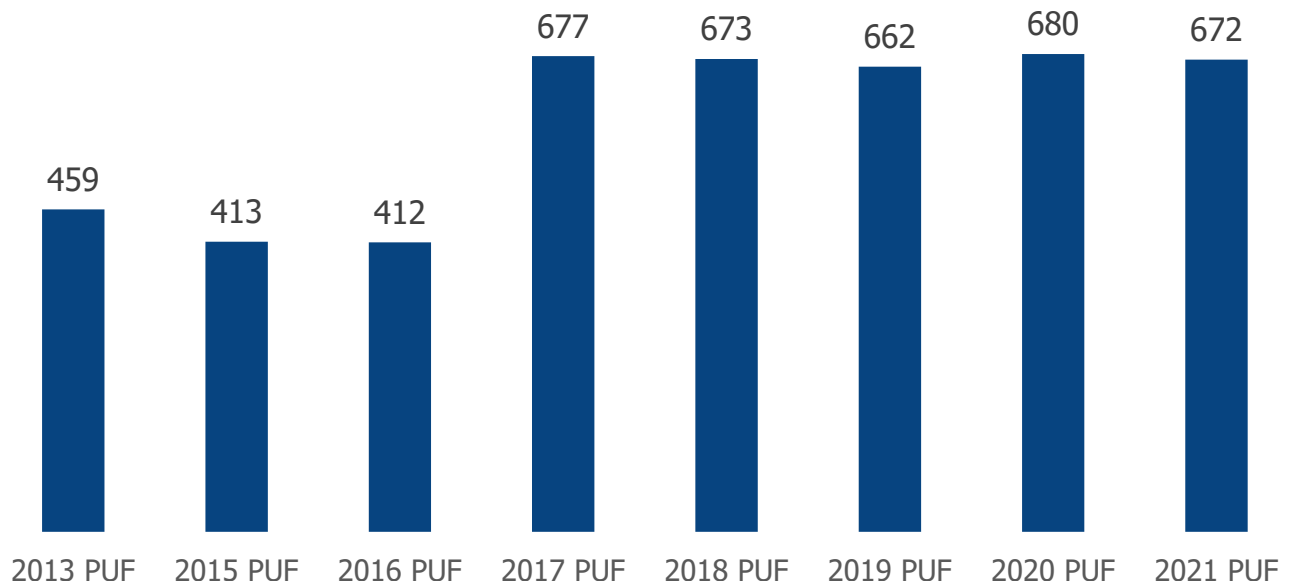
Exhibit 1.1.2 presents a summary of the content updates in the 2021 MCBS Survey File PUF.

**Exhibit 1.1.2:** 2021 MCBS Survey File PUF Content Updates

| Segment | Variable | Variable Label | Description |
|---|---|---|---|
| Fall | HLT_OCOSARTH | Osteoarthritis (ever) | Added |
| Fall | HLT_OCARTHOT | Arthritis, other (ever) | Added |
| Fall | ACC_COLAGNCY | Collection agency due to medical bills | Dropped |
| Fall | HOU_D_COMPHH | Household Composition Code | Dropped |
| Fall | PRV_COLNRCNT | How long since colonoscopy/sigmoid | Dropped |
| Fall | PRV_OCCEXAM | Most recent oral/mouth cancer exam | Dropped |
| Winter | KNW_COMPDESK | Own/use a desktop/laptop computer | Added |
| Winter | KNW_COMPPHON | Own/use a smartphone | Added |
| Winter | KNW_COMPTAB | Own/use a tablet/portable computer | Added |
| Winter | TMW_TELMED | Usual provider offers telehealth | Added |
| Winter | TMW_TELMEDT1 | Provider offers telephone or video appts | Added |
| Winter | TMW_TELMEDDU | Past yr, provider offered telehealth appt | Added |
| Winter | TMW_TELMEDT3 | Offered appt was telephone or video | Added |
| Winter | TMW_TELMEDUS | Past yr, had telehealth appt | Added |
| Winter | TMW_TELMEDT4 | Appt was telephone or video | Added |
| Winter | KNW_KNHAVCOM | Have a personal computer | Dropped |
| Winter | ACW_D_MDVIST | MD visit in this or any of prev 2 rounds | Dropped |
| Winter | ACW_MDMCOND | Reason Dr seen: med condition | Dropped |
| Winter | ACW_MDTESTS | Reason Dr seen: tests | Dropped |
| Winter | ACW_MDFOLUP | Reason Dr seen: follow-up | Dropped |
| Winter | ACW_MDCHKUP | Reason Dr seen: checkup | Dropped |
| Winter | ACW_MDRFRL | Reason Dr seen: referral | Dropped |
| Winter | ACW_MDSURGY | Reason Dr seen: surgery | Dropped |
| Winter | ACW_MDPSHOT | Reason Dr seen: preventative shot | Dropped |
| Winter | ACW_MDTSHOT | Reason Dr seen: treatment shot | Dropped |
| Winter | ACW_MDPMED | Reason Dr seen: medication | Dropped |
| Winter | ACW_MDOTHER | Reason Dr seen: other | Dropped |
| Winter | ACW_MDAPPT | Was visit to Dr appointment or walk-in | Dropped |
| Winter | ACW_MDDRTEL | Dr's office tell come back | Dropped |
| Winter | ACW_D_MDAPPT | Days waited for Dr's appointment | Dropped |
| Winter | ACW_D_MDTIME | Minutes wait at Dr | Dropped |
| Summer | CVS_VCNUMS | Number of COVID-19 vaccine doses reported by Summer 2022 | Added |

# 2. OVERVIEW OF THE MCBS

Medicare is the nation's health insurance program for persons 65 years and over and for persons younger than 65 years who have a qualifying disability. The MCBS is sponsored by CMS and contains data provided by a representative national sample of the Medicare population. The MCBS is designed to aid CMS in administering, monitoring, and evaluating the Medicare program. A leading source of information on Medicare and its impact on beneficiaries, the MCBS provides important information on beneficiaries that is not otherwise collected through operational or administrative data from the Medicare program and plays an essential role in the monitoring and evaluation of beneficiary health status and health care policy.

The MCBS is a continuous, multi-purpose longitudinal survey, representing the population of beneficiaries aged 65 and over and beneficiaries aged below 65 with certain disabling conditions, residing in the United States. Most interviews were traditionally conducted in-person in households and facilities using computer-assisted personal interviewing (CAPI). However, due to the COVID-19 pandemic, data collection switched to phone-only interviews in March 2020 and throughout most of 2021 with a gradual return to some in-person interviewing beginning in November 2021. MCBS data collection will include both in-person and phone interviewing going forward. Fieldwork for the first round of data collection began in September 1991; since then, the MCBS has continued to collect and provide essential data on the costs, use, and health care status of Medicare beneficiaries. The MCBS has been continuously conducted since 1991, completing more than 1.2 million interviews provided by thousands of respondents.

The MCBS primarily focuses on economic and beneficiary topics including health care use and health care access barriers, health care expenditures, and factors that affect health care utilization. As a part of this focus, the MCBS collects a variety of information about the beneficiary, including demographic characteristics, health status and functioning, access to care, insurance coverage and out of pocket expenses, financial resources, and potential family support. The MCBS collects this information in three data collection periods, or rounds, per year. Over the years, data from the MCBS have been used to inform many advancements to the Medicare program, including the creation of new benefits such as Medicare's Part D prescription drug benefit.

This *Data User's Guide* uses the following definitions for beneficiary and respondent:

- Beneficiary refers to a person receiving Medicare services who may or may not be participating in the MCBS. Beneficiary may also refer to an individual selected from the MCBS sample about whom the MCBS collects information.[3]

- Respondent refers to a person who answers questions for the MCBS; for Community interviews, this person can be the beneficiary or a proxy.

For questions or suggestions on this document or other MCBS data-related questions, please email MCBS@cms.hhs.gov.

---

[3] https://www.cms.gov/Medicare/Medicare-General-Information/MedicareGenInfo/index.html.

# 3. TECHNICAL AND PROGRAMMING INFORMATION

## 3.1 General Information

The 2021 MCBS Survey File PUF is divided into three segments (Fall, Winter, and Summer), which allows for the release of data collected in all rounds. The Fall segment includes data for 12,783 sampled beneficiaries. The samples for the Winter and Summer segments are subsets of the Fall segment sample. The Winter segment includes data for 10,628 sampled beneficiaries, and the Summer segment includes data for 7,538 sampled beneficiaries. All three segments include survey weights that allow for analysis that is nationally representative of the population of beneficiaries ever enrolled in Medicare at any point in 2021.

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 three 2021 MCBS Survey File PUF segments only and cannot be used for linking to MCBS data files other than the MCBS Survey File PUF. The PUF_ID does link a beneficiary between the three 2021 MCBS Survey File PUF segments. Each beneficiary's PUF_ID is randomly generated each year, so it is not possible to link a beneficiary's data between years, and the value of the PUF_ID does not provide any information about the beneficiary.

All variables in the MCBS Survey File PUF are in numeric or integer formats. Formats and values for each variable are available in the MCBS Survey File PUF codebook.

Variable groups contain prefixes to help users identify these groups by topic area. Exhibits 3.1.1 – 3.1.3 include information about these variable prefixes and the locations of the corresponding variables in the Survey File LDS data segments (i.e., the 2021 LDS Survey File individual files).

**Exhibit 3.1.1:** 2021 MCBS Survey File PUF Fall Segment Variable Prefixes, Number of Variables, Descriptions, and Related LDS Survey File Segments

| MCBS Survey File PUF Variable Prefix | Description | Number of PUF Variables in Grouping | LDS Survey File Data Segments |
|---|---|---|---|
| ADM_ | Administrative data | 22 | ADMNUTLS, HISUMRY, HITLIN |
| INS_ | Insurance status, coverage, and type | 11 | HISUMRY, HITLINE |
| INT_ | Interview characteristics | 2 | INTERV |
| DEM_ | Age, sex, and race groups | 11 | DEMO |
| ACC_ | Access to, use, and satisfaction with health care | 18 | ACCESSCR, SATWCARE, ASSIST |
| HLT_ | Health conditions and limitations in activities of daily living | 88 | GENHLTH, VISHEAR, NAGIDIS, CHRNCOND |
| PRV_ | Preventive care and physical activity | 26 | PREVCARE, NAGIDIS |
| RSK_ | Health behavior risk factors | 6 | NICOALCO |
| FAL_ | Falls | 12 | FALLS |
| HOU_ | Housing characteristics | 29 | HHCHAR |
| MA_ | Medicare Advantage supplement | 6 | MAPLANQX |
| PUFF | PUF ever enrolled weights (fall) | 101 | EVRWGTS |

**Exhibit 3.1.2:** 2021 MCBS Survey File PUF Winter Segment Variable Prefixes, Number of Variables, Descriptions, and Related LDS Survey File Segments

| MCBS Survey File PUF Variable Prefix | Description | Number of PUF Variables in Grouping | LDS Survey File Data Segments |
|---|---|---|---|
| ACW_ | Access to care (winter round) | 71 | ACCSSMED, USCARE |
| KNW_ | Medicare knowledge | 14 | MCREPLNQ |
| TMW_ | Telehealth services | 6 | TELEMED |
| PUFW | PUF ever enrolled weights (winter) | 101 | N/A |

**Exhibit 3.1.3:** 2020 MCBS Survey File PUF Summer Segment Variable Prefixes, Number of Variables, Descriptions, and Related LDS Survey File Segments

| MCBS Survey File PUF Variable Prefix | Description | Number of PUF Variables in Grouping | LDS Survey File Data Segments |
|---|---|---|---|
| CVS_ | COVID-19 vaccine uptake | 3 | COMMDOSE |
| FIS_ | Food insecurity | 4 | FOODINS |
| RXS_ | Prescription medication use and access | 29 | RXMED |
| PUFS | PUF ever enrolled weights (summer) | 101 | N/A |

## 3.2 Data File Information

Detailed information about variables in the MCBS Survey File PUF can be found in the PUF codebooks. The codebook includes SAS® variable names, labels, a note to indicate which beneficiaries were eligible for the question, the question number for the question that was asked in the survey, and a label which summarizes the question text. Certain variables in the PUF were recoded due to disclosure concerns so the categories in the PUF codebook may differ from the categories in the Questionnaires and Questionnaire User Documentation (e.g., "no usable vision" for variable HLT_ECTROUB and "deaf" for HLT_HCTROUB both reflect such recoding). Other variables were created by combining two variables, and their variable label indicates a recoded variable (e.g., HLT_ALZDEM).

For each variable, the formats and format values are included in the codebook:

- Values of .R indicate "refused" and .D indicate "don't know."
- All values of "inapplicable" have been combined with missing values.
- Unweighted frequencies of most variables included in the MCBS Survey File PUF are provided in the accompanying codebook file.

The MCBS Survey File PUF datasets are saved as SAS export files. Directions and sample SAS code are given below and also in Appendix B to help users read the datasets into SAS.

Assume the MCBS 2021 Survey File PUF export files (e.g., PUF20YY_1_FALL.xpt) are downloaded into the folder "C:\MCBS\DOWNLOAD". The "YY" in "PUF20YY" refers to the data year of the MCBS Survey File PUF. The following SAS code can then be used to import the Survey File PUF Fall segment into SAS:

```
LIBNAME PUFLIB 'C:\MCBS\SASDATA';
FILENAME F "C:\MCBS\DOWNLOAD\PUF20YY_1_FALL.XPT";
PROC CIMPORT LIBRARY=PUFLIB INFILE=F;
```

RUN;

Likewise, the SAS code above could be altered to import the Survey File PUF Winter segment (SAS export file PUF20YY_2_WINTER.xpt) or Survey File PUF Summer segment (SAS export file PUF20YY_3_SUMMER.xpt) into SAS. Additionally, a comma-separated values (CSV) file is available for use with other statistical software packages such as R® and STATA®.

A text file with SAS programming code to import the .xpt files, create formats, and apply SAS labels is provided for users.

## 3.3 Comparison to the LDS

The MCBS Survey File PUF differs from the MCBS Survey File LDS, because it has been evaluated for disclosure risk and additional steps were taken to protect beneficiary confidentiality. The 2021 MCBS Survey File PUF contains data for 12,7836 sampled beneficiaries and 672 variables, which is similar to the number of beneficiaries contained in the community-only 2021 MCBS Survey File LDS segments, but with fewer variables. Many Survey File LDS variables that posed a disclosure risk were dropped or recoded to create the variable set for the MCBS Survey File PUF.

Due to disclosure concerns, the MCBS Survey File PUF includes only beneficiaries living in the community for at least one interview during the data collection year, and thus excludes all beneficiaries who were in a health care facility during all interviews that year (n=1,560). Variables that were only created for facility residents are excluded.[4] Additionally, the MCBS Survey File PUF contains no health care utilization, cost, or payment data (including Medicare claims data) for individual beneficiaries.

The MCBS Survey File PUF is free and available for download on the CMS website. For users interested in the MCBS Survey File and Cost Supplement File LDS, more information on the LDS process can be found at: https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS.

A summary of the differences between the two data products is presented in Exhibits 3.3.1.a and 3.3.1.b.

---

[4] 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.

**Exhibit 3.3.1.a:** Comparison between the 2021 MCBS Survey File PUF and 2021 MCBS Survey File LDS
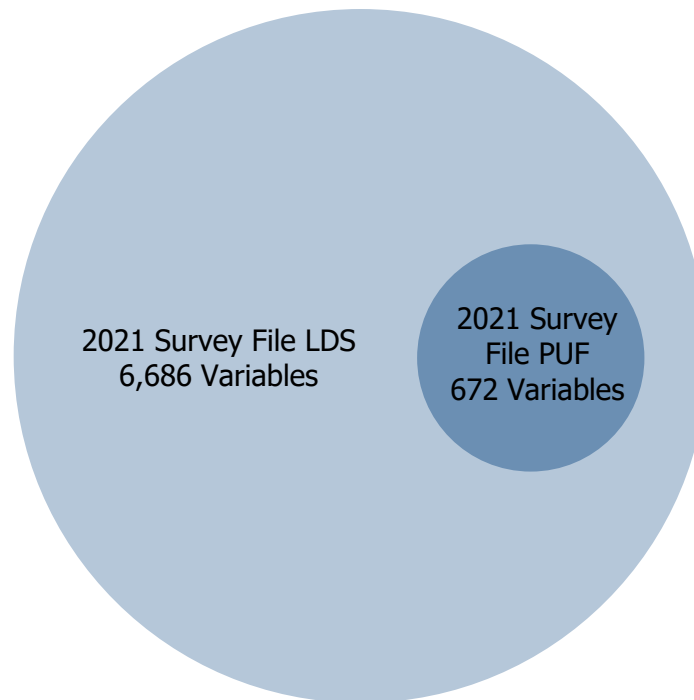


**Exhibit 3.3.1.b:** Comparison between the 2021 MCBS Survey File PUF and 2021 MCBS Survey File LDS

| Domain | MCBS Survey File PUF | MCBS Survey File LDS |
|---|---|---|
| **Population** | Community | Community and facility |
| **Number of variables** | 672 variables across three data segments | Over 6,000 variables across 46 data segments |
| **ID** | PUF_ID; Randomly generated, can't be linked back to BASEID, changes each year | BASEID; Randomly generated, can't be linked back to Medicare Beneficiary Identifier (MBI), consistent between years |
| **Date fields** | NO | YES |
| **Geographic identifiers** | NO | YES |
| **Cost/payment data** | NO | YES |
| **Demographic data** | YES; All variables are categorical, limited age categories | YES; Continuous, all age variables available |
| **Insurance coverage** | YES; Summarized to annual level | YES; Monthly level |
| **Identifiable plan-related information for MA or Part D** | NO | YES |
| **Population weights** | Ever enrolled weights only | Both ever enrolled and continuously enrolled weights |

Exhibit 3.3.2 details the socio-demographic variables and values available in the MCBS Survey File PUF and the MCBS Survey File LDS, for comparison purposes. Variables in the MCBS Survey File LDS, which do not have an equivalent in the MCBS Survey File PUF, are noted below the exhibit.[5]

**Exhibit 3.3.2:** Comparison of Socio-demographic Variables and Values in the MCBS Survey File PUF and MCBS Survey File LDS

| Socio-demographic Characteristic | MCBS Survey File PUF Variables and Values | MCBS Survey File LDS Variables and Values |
|---|---|---|
| **Gender** | DEM_SEX (Gender): Male; Female | ROSTSEX (Gender): Male; Female |
| **Age** | DEM_AGE (Age group): <65 years; 65-74 years; ≥75 years | D_STRAT (MCBS Sample age stratum): 0-44 years; 45-64 years; 65-69 years; 70-74 years; 75-79 years; 80-84 years; ≥ 85 years<br><br>H_AGE (Age of beneficiary): Age of beneficiary in years |
| **Armed Services** | DEM_ARMED (Ever served in armed forces (AF)): Yes; No | SPAFEVER (Ever served in AF): Yes; No<br><br>SPAFVIET: Served in AF during Vietnam era; SPAFKORE: Served in AF during Korean conflict; SPAFWWII: Served in AF during WWII; SPAFGULF: Served in AF during Gulf War; SPAFIRAF: Served in AF during Iraq/Afghanistan conflict; SPAFPEAC: Served in AF during peace time; SPNGEVER: Ever active in National Guard/Reserve; SPNGALL: All active duty spent in National Guard<br><br>SPNGDSBL: Disability from service<br><br>SPAVARATE: Current VA disability rating |
| **Education** | DEM_EDU (Highest grade completed): Less than high school; High school or vocational, technical, business, etc.; More than high school | SPDEGRCV (Highest grade sampled person (SP) completed): No schooling; Nursery to 8th grade; 9th to 12th grade, but no diploma; High school graduate; Vocational, technical, business, etc.; Some college, but no degree; Associate's degree; Bachelor's degree; Graduate or professional degree |
| **Income** | DEM_INCOME (Income group of SP and Spouse): <$25,000; ≥$25,000 | INCOME (Income range of SP and spouse): <$5,000; $5,000 - $9,999; $10,000 - $14,999; $15,000 - $19,999; $20,000 - $24,999; $25,000 - $29,999; $30,000 - $39,999; $40,000 - $49,999; $50,000 - $59,999; $60,000 - $79,999; $80,000 - $99,999; $100,000 - $119,999; $120,000 - $139,999; ≥$140,000<br><br>INCOME_H (SP and spouse total income last year): Range of values |

[5] The MCBS Survey File LDS contains additional socio-demographic information, including location of residence, rural-urban commuting area details, the number of living children the beneficiary has, employment status, status of Social Security Administration (SSA) check, and English proficiency, which do not have corresponding variables available in the MCBS Survey File PUF. Please note that additional race/ethnicity variables from administrative sources are included in the MCBS Survey File LDS.

| Socio-demographic Characteristic | MCBS Survey File PUF Variables and Values | MCBS Survey File LDS Variables and Values |
|---|---|---|
| **Interview Language or Language Spoken at Home** | INT_LANG (Language of interview): English; Spanish | INTLANG (Language of interview): English; Spanish<br><br>WHATLANG (Language spoken at home): Spanish; French; German; Italian; Tagalog; Chinese; Russian; Greek; Filipino; Portuguese; Hindi; Other<br><br>OTHRLANG (Language other than English spoken at home): Yes; No |
| **Marital Status** | DEM_MARSTA (Marital status): Married; Widowed; Divorced/separated; Never married | SPMARSTA (Marital status of SP): Married; Widowed; Divorced; Separated; Never married |
| **Metro Status** | DEM_CBSA (Metro status): Metro area; Non-metro area | H_CBSA (Type of CBSA as designated by CBSA): Metropolitan area-population of ≥50,000; Micropolitan area-population between 10,000 to 50,000; Non-CBSA |
| **Race/Ethnicity** | DEM_RACE (Race/ethnicity group): Non-Hispanic White; Non-Hispanic Black; Hispanic; Other | D_RACE2 (Race of SP): Asian; African American; Native Hawaiian or Pacific Islander; White; American Indian or Alaska Native; More than one<br><br>RACEAS: Asian; RACEASAI: Asian Indian; RACEASCH: Chinese; RACEASFI: Filipino; RACEASJA: Japanese; RACEASKO: Korean; RACEASVI: Vietnamese; RACEASOT: Other Asian; RACEAA: Black or African-American; RACENH: Native Hawaiian or Pacific Islander; RACEPIHA: Native Hawaiian; RACEPIGU: Guamanian Chamorro; RACEPISA: Samoan; RACEPIOT: Other Pacific Islander; RACEWH: Caucasian; RACEAI: American Indian or Alaska Native<br><br>HISPORIG (Is SP of Hispanic or Latino origin?): Yes; No<br><br>HISPORMA: Mexican/Mex American/Chicano; HISPORPR: Puerto Rican; HISPORCU: Cuban; HISPOROT: Other Hispanic/Latino/Spanish origin |

| Socio-demographic Characteristic | MCBS Survey File PUF Variables and Values | MCBS Survey File LDS Variables and Values |
|---|---|---|
| **Race/Ethnicity by Age Group** | DEM_RE_AGE (Race/ethnicity age group): Non-Hispanic White, <65 years; Non-Hispanic White, 65-74 years; Non-Hispanic White, 75-84 years; Non-Hispanic White, 85+ years; Non-Hispanic Black, <65 years; Non-Hispanic Black, 65-74 years; Non-Hispanic Black, 75-84 years; Non-Hispanic Black, 85+ years; Hispanic, <65 years; Hispanic, 65-74 years; Hispanic, 75+ years; Other, <65 years; Other, 65-74 years; Other, 75+ years | Note: An MCBS Survey File LDS user could construct a similar race/ethnicity by age variable using D_RACE2, HISPORIG, and D_STRAT. |
| **Income Poverty Ratio Medicare Threshold** | DEM_IPR_IND (Income Poverty Ratio Medicare Threshold): <=100% of the Federal Poverty Level; >100% and <=120% of the Federal Poverty Level; >120% and <=135% of the Federal Poverty Level; >135% and <=200% of the Federal Poverty Level; >200% of the Federal Poverty Level | IPR (Income Poverty Ratio Medicare Threshold): <=100% of the Federal Poverty Level; >100% and <=120% of the Federal Poverty Level; >120% and <=135% of the Federal Poverty Level; >135% and <=200% of the Federal Poverty Level; >200% of the Federal Poverty Level |

# 4. SURVEY OVERVIEW

## 4.1 Design of MCBS

In its initial design, the MCBS was to serve as a traditional longitudinal survey of the Medicare population. There was no predetermined limit to the duration of time a beneficiary, once selected to participate, was to remain in the sample. However, beginning in 1994, participation of beneficiaries in the MCBS was limited to no more than four years.

Although participation in the survey is limited to four years, MCBS data collection is continuous throughout the year with three distinct seasons (i.e., rounds) of data collection per year. In general, the three rounds are: winter (January through April); summer (May through August); and fall (September through December). The primary reason for the round to round design is to create shorter recall periods during the year to capture more complete and accurate health care costs and utilization for beneficiaries.

The 2021 MCBS data releases reflect data collected from January 2021 through December 2021 (see Exhibit 4.1.1), as well as Topical sections, income and assets data, and chronic pain information collected through the Winter and Summer 2022 rounds.[6] Exhibit 4.1.1 depicts an MCBS data collection year and the typical span of the rounds.

**Exhibit 4.1.1:**  Typical MCBS Data Collection Year
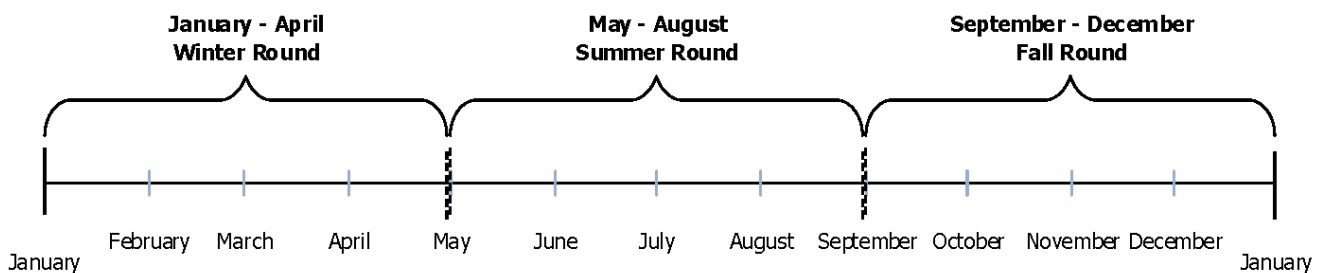


The initial interview of newly-selected beneficiaries takes place in the fall round. Since 2016, the fall round begins early (in late July or early August) to allow more time to locate and conduct outreach to the new survey respondents.

Subsequent rounds, which occur every four months, involve the re-interviewing of the same beneficiary (or appropriate proxy respondents or facility staff) over a four year period (up to 11 interviews in total). Exhibit 4.1.2 depicts the timeline of participation for beneficiaries selected to be in the MCBS sample.

---

[6] Due to the nature of some survey items, PUF data for each data year may include data pulled forward from a prior data collection year and/or data added from a future data collection year due to the specific reference period.

**Exhibit 4.1.2:**  MCBS Beneficiary Participation Timeline
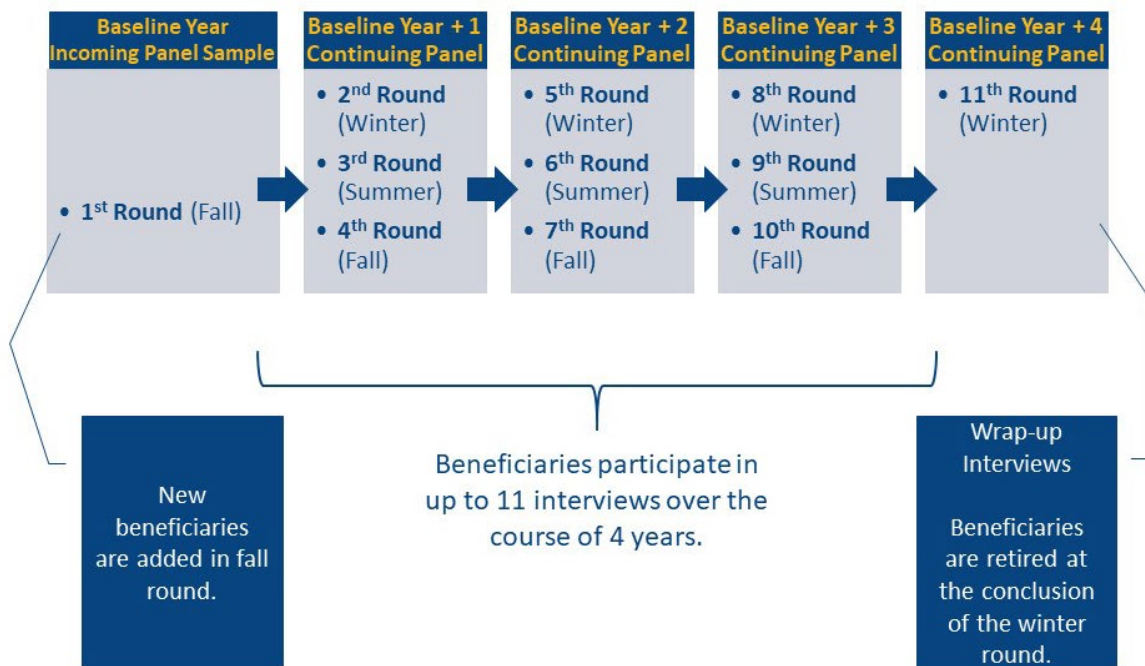


## 4.2 Sample Design

The MCBS uses a rotating panel sample design, covering the population of Medicare beneficiaries residing in the continental U.S. (48 states and the District of Columbia) for the survey year.[7] Each MCBS panel, an annual statistical sample of all Medicare enrollees, is interviewed up to three times a year over a four-year period creating a continuous profile of selected beneficiaries' health care experiences.[8] One panel is retired at the conclusion of each winter round, and a new panel is selected to replace it each fall round (see Exhibit 4.2.1). The size of the new panel is designed to provide a stable number of beneficiaries across all panels participating in the survey annually.

---

[7] Alaska and Hawaii are not included among the states from which the sample is selected due to the high cost of data collection in those areas; however, they are included in control totals for weighting purposes. Beginning in 2017, sampling from Puerto Rico was discontinued. Beginning in 2018, all data collection in Puerto Rico was discontinued.

[8] The three rounds per year are referred to seasonally. Respondents are interviewed in the winter round, the summer round, and the fall round each year.

**Exhibit 4.2.1:** 2018-2021 MCBS Rotating Panel Design

| Data Collection Schedule | | | Panel | | | |
|---|---|---|---|---|---|---|
| Calendar Year | Season | Round# | 2018 | 2019 | 2020 | 2021 |
| **2018** | Winter | 80 | | | | |
| | Summer | 81 | | | | |
| | Fall | 82 | | | | |
| **2019** | Winter | 83 | | | | |
| | Summer | 84 | | | | |
| | Fall | 85 | | | | |
| **2020** | Winter | 86 | | | | |
| | Summer | 87 | | | | |
| | Fall | 88 | | | | |
| **2021** | Winter | 89 | | | | |
| | Summer | 90 | | | | |
| | Fall | 91 | | | | |

The MCBS employs a three-stage cluster sample design. Primary sampling units (PSUs) are made up of major geographic areas consisting of metropolitan areas or groups of rural counties. Secondary sampling units (SSUs) are made up of census tracts or groups of tracts within the selected PSUs. Medicare beneficiaries, the ultimate sampling units (USUs), are then selected from within the selected SSUs. The final 2021 MCBS Panel was drawn from 104 PSUs, which contained 685 SSUs. The MCBS sample is annually "supplemented" during the fall round to account for attrition (deaths, dis-enrollments, refusals) and current-year enrollees. Each annual supplement is referred to as the Incoming Panel sample.

Beneficiaries for the MCBS are sampled from the Medicare Administrative enrollment data. The beneficiaries included in the MCBS Survey File PUF represent a randomly selected cross-section of all beneficiaries who were ever enrolled in either Part A or Part B of the Medicare program for any portion of 2021.[9] The MCBS Survey File PUF represents four separate MCBS panels identified by the year in which the panel was selected and first interviewed (i.e., for the 2021 MCBS Survey File PUF, the 2018, 2019, 2020, and 2021 Panels). Exhibit 4.2.2 shows the distribution of each of the four panels included in the 2021 MCBS Survey File PUF.

For more information on the sample design, please see the *Survey File Data User's Guide* at: https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS/Codebooks.

**Exhibit 4.2.2:** 2021 MCBS Composition of Panels in the MCBS Survey File PUF

| Data Year (Fall) | Number of Beneficiaries Selected |
|---|---|
| 2018 | 1,914 |
| 2019 | 2,211 |
| 2020 | 3,178 |
| 2021 | 5,480 |

---

[9] While beneficiaries included in the LDS releases represent both the ever enrolled and continuously enrolled Medicare population, the MCBS Survey File PUF solely represents the ever enrolled population.

## 4.3 Eligibility

### 4.3.1 Medicare Population Covered by the 2021 LDS and MCBS Survey File PUF

Beneficiaries who became eligible for Medicare Part A or B and enrolled anytime during the sampling year were eligible to be sampled as part of the annual panel.

## 4.4 Case Types

MCBS beneficiaries are classified by their phase of survey participation (i.e., Incoming or Continuing) and interview participation (i.e., Community or Facility), which is determined by residence status. Although they appear in the MCBS LDS releases, beneficiaries for whom only Facility interviews were conducted during the data collection period are not included in the MCBS Survey File PUF. Researchers interested in the population of beneficiaries living in facilities will need to use the MCBS LDS, as discussed in Section 3.3.

### 4.4.1 Incoming and Continuing Cases

Every fall round of data collection, a new panel of sampled beneficiaries is added to the total sample to replace the panel of sampled beneficiaries completing a final interview and exiting the MCBS in the prior winter round. Newly selected MCBS beneficiaries begin in the fall round and are referred to as Incoming Panel cases. After the initial interview, they are referred to as Continuing cases.

### 4.4.2 Community Interviews

Approximately 93 percent of the interviews are held with the beneficiaries or proxies who are living in their own residence or with family or friends. These interviews are called Community interviews; the remaining 7 percent of the interviews are administered for beneficiaries living in a facility, and these beneficiaries are not included in the MCBS Survey File PUF.

Over the course of a four-year period, however, it is not uncommon for beneficiaries to enter long-term care facilities (e.g., nursing homes) or to go back and forth between community and facility settings. In order to obtain an accurate representation of the experiences of all Medicare beneficiaries, the MCBS includes beneficiaries wherever they reside, even if they enter or reside in a facility for the duration of their four years with the study. The MCBS Survey File PUF excludes those beneficiaries who were in a facility for each interview, due to disclosure concerns.

## 4.5 Interviewing and Training Procedures

### 4.5.1 Overview of Data Collection

CMS contracts with NORC at the University of Chicago (NORC) to administer the MCBS. A national team of specially trained and certified NORC field interviewers conduct either interviews with MCBS beneficiaries or their designated proxies or they conduct interviews with Facility administrators on behalf of beneficiaries. The first interview conducted for an Incoming Panel beneficiary is relatively short, as it does not collect health care utilization or cost data. Continuing interviews are longer, as field interviewers collect information about the beneficiary's health care utilization and associated costs.

Due to the COVID-19 pandemic, data collection was exclusively conducted by phone from January 2021 through October 2021, when in-person visits to gain cooperation and interviews were permitted in local areas where COVID-19 case data met acceptable county-level thresholds. MCBS interviewer staff who volunteered

and were fully vaccinated for COVID-19 were eligible to participate with in-person data collection; 93 interviewers were trained and certified to conduct in-person work during Fall 2021. Initial in-person activities focused on following up with Incoming Panel beneficiaries who had been difficult to reach through mail and phone outreach alone. These gaining cooperation visits focused on connecting with beneficiaries, sharing materials about the survey, and setting appointments for interviews. Field staff were directed to focus in-person interviewing on cases with very high health care utilization, along with cases where the respondent has difficulty seeing and/or hearing. Both factors result in more burdensome phone interviewing.

### 4.5.1.1 Overview of Recruitment of Beneficiaries and Scheduling Procedures

Medicare beneficiaries selected to participate in the MCBS receive a letter and a brochure in the mail, introducing the study and explaining that an interviewer from NORC will contact them to schedule an appointment. For Incoming Panel interviews, initial contact is typically made in person; for Continuing interviews, outreach to set an appointment is most often made by phone. If beneficiaries are unable to answer questions or require language assistance, they can enlist the help of an assistant, such as a family member, to help complete the interview; a proxy can also respond on behalf of the beneficiary if the beneficiary is incapacitated or unable to complete the interview. For Spanish speaking respondents, a Spanish version of the Community Questionnaire is available, and bilingual interviewers conduct the interview.

### 4.5.1.2 Computer-Assisted Personal Interviewing (CAPI)

Field interviewers complete MCBS interviews using a CAPI instrument loaded on a laptop. The CAPI program automatically guides the field interviewer through the questions, records the answers, and contains logic and skip flows that increase the output of timely and high quality data. The CAPI also contains follow-up questions where data were missing from the previous interview. When the interview is completed, the CAPI system allows the field interviewer to transmit the data electronically to the NORC central office in a secure manner.

## *4.5.2 Interviewer Training*

Nationally, the MCBS employs an average of approximately 180 field interviewers,[10] who participate in a combination of several targeted training initiatives and careful coaching and monitoring activities throughout data collection.

The 2021 MCBS Training Program consisted of remote trainings which varied based on the level of experience of the interviewer (new to MCBS or MCBS-experienced), the type of interview component (Community or Facility), the sample type (incoming Panel or Continuing), and the season-specific requirements (new or changing questionnaire sections or data collection protocols). No in-person trainings were held due to restrictions and risk associated with the COVID-19 public health emergency. The program was structured to expose all field staff to the same training content, ensuring that the performance of data collection responsibilities was standardized, methodical, and measurable.

In an effort to meet all interviewers' skill-building and training needs, NORC continued to work with field managers to ensure interviewers receive training during each data collection round, such as weekly field memos, interviewer group call sessions, and interviewer observations via "call-alongs." For most interviewers, this included an interview observation by a mentor via conference call and screen-sharing with an experienced mentor. These methods covered important data collection tips, provided answers to interviewer questions, and offered reminders about how to handle complicated scenarios, especially via phone interviewing.

---

[10] The fall round starts with a higher number of field interviewers which, over the course of the year, is reduced due to staff turnover. Each summer, a small cohort of new interviewers is hired for the MCBS.

## 4.5.3 Privacy and Data Security

Field interviewer training stresses the importance of maintaining privacy, and project protocols are documented within the field interviewer manual. Field outreach and contacting procedures also maintain and ensure confidentiality. These procedures include the utilization of standard computer security protocol (dual authentication password protection for each interviewer laptop) and restrictions on submitting personally identifiable information (PII) through electronic mail. All MCBS survey staff directly involved in data collection and/or analysis activities are required to sign a Non-Disclosure Agreement and a confidentiality agreement.

NORC and CMS are committed to protecting respondent confidentiality and privacy, and both organizations diligently uphold provisions established under the Privacy Act of 1974, the NORC Institutional Review Board (IRB), the Office of Management and Budget (OMB), and the Federal Information Security Management Act of 2002. As stated in the MCBS OMB documentation, the information collected for MCBS is protected by NORC and by CMS. Respondent data are used only for research and statistical purposes. As required under the Privacy Act of 1974, identifiable information is not disclosed or released without the consent of the individual or the establishment, except to those involved in research (Public Law 93-579). The MCBS is authorized by section 1875 (42 USC 139511) of the Social Security Act and is conducted by NORC at the University of Chicago for the U.S. Department of Health and Human Services. The OMB Number for this survey is 0938-0568.

# 5. QUESTIONNAIRES

## 5.1 Overview

The MCBS Questionnaire structure features two components (Community and Facility), administered based on the beneficiary's residence status. Within each component, the flow and content of the questionnaire varies by interview type and data collection season (fall, winter, or summer). There are two types of interviews (Baseline and Continuing) containing two types of questionnaire sections (Core and Topical). See Exhibit 5.1 within the *Survey File Data User's Guide* for a depiction of the MCBS Questionnaire structure: https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS/Codebooks.

- Community Component: Survey administered for beneficiaries living in the community (i.e., not in a long-term care facility such as a nursing home) during the reference period covered by the MCBS interview. An interview may be conducted with the beneficiary or a proxy.

- Facility Component: Survey administered for beneficiaries living in facilities, such as long-term care nursing homes or other institutions, during the reference period covered by the MCBS interview. Interviewers conduct the Facility component with staff members located at the facility (i.e., Facility respondents); beneficiaries are not interviewed if they reside at a facility.

Within each component, there are two types of interviews – a Baseline interview and a Continuing interview.

- Baseline: The initial questionnaire administered in the fall round of the year the beneficiary is selected into the sample (interview #1).

- Continuing: The questionnaire administered as beneficiaries progress through the study (interviews #2-11).

Depending on the interview type and data collection season (fall, winter, or summer), the MCBS Questionnaire includes Core and Topical sections:

- Core: These sections collect data central to policy goals of CMS. Core sections collect information related to socio-demographics, health insurance coverage, health care utilization and costs, beneficiary health status, and experiences with care, as well as operational and procedural data. The questionnaire sections may be asked each round or seasonally (fall, winter, summer).

- Topical: These sections collect information on special interest topics. They may be fielded every round or on a seasonal basis. Specific topics may include housing characteristics, health behaviors, knowledge about Medicare, and health-related decision making.

See Sections 5.2 and 5.3 within the *Survey File Data User's Guide* for additional detail on the 2021 Core and Topical sections: https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS/Codebooks.

# 6. SAMPLING

## 6.1 Medicare Population Covered by the 2021 MCBS Survey File PUF

The MCBS data releases are a reflection of enrolled Medicare beneficiaries residing in the continental United States. The sample for the MCBS is drawn from a subset of the Medicare enrollment data, which is a list of all Medicare beneficiaries. Excluded are residents of foreign countries and U.S. possessions and territories. The MCBS Survey File PUF further excludes Medicare beneficiaries who only provided facility-based interviews during the data year. The MCBS data releases include two overlapping but differing populations:

- The ever enrolled population represents individuals who were enrolled in Medicare at any time during the calendar year. This population includes beneficiaries who enrolled during the calendar year 2021 as well as those who dis-enrolled or died prior to their fall interview.[11] The ever enrolled population includes beneficiaries who were enrolled in Medicare for at least one day at any point during 2021.

- The continuously enrolled population represents only those individuals continuously enrolled in Medicare from January 1, 2021 up to and including their fall interview; this specifically excludes beneficiaries who enrolled during the calendar year 2021 and those who dis-enrolled or died prior to their fall interview. The concept of continuously enrolled is consistent with the concept of being exposed or "at risk" for using services up to and including their fall interview.

The MCBS Survey File PUF includes weights that represent the ever enrolled population, whereas the MCBS LDS releases include separate sets of weights that represent both the ever enrolled and continuously enrolled Medicare population.

Exhibits 6.1.1 and 6.1.2 present estimates of the size of the ever enrolled Medicare population living in the community by race and age (as of December 31, 2021), by sex in the 2021 MCBS Survey File PUF. Exhibit 6.1.3 presents the aggregated estimates of the size of the ever enrolled Medicare population living in the community overall and by sex and race.

**Exhibit 6.1.1:** Estimated Number of Male Community Medicare Beneficiaries by Race and Age, in the 2021 MCBS Survey File PUF*

| Race | Age as of 12/31/2021 | Weighted Count |
|---|---|---|
| White non-Hispanic | Under 65 years | 2,370,332 |
| | 65-74 years | 11,104,942 |
| | 75+ years | 7,416,766 |
| Black non-Hispanic | Under 65 years | 738,572 |
| | 65-74 years | 1,410,029 |
| | 75+ years | 553,055 |
| Hispanic | Under 65 years | 461,192 |
| | 65-74 years | 1,090,411 |
| | 75+ years | 624,435 |
| Other† | Under 65 years | 275,062 |
| | 65-74 years | 1,039,354 |
| | 75+ years | 539,411 |

---

[11] Note that data collection for beneficiaries who enrolled during 2021 and died in 2021 after enrollment but before their fall interview was still pursued through attempts at conducting proxy interviews.

SOURCE: 2021 MCBS Survey File PUF, weighted counts.

*Weighted counts may not sum to the total of beneficiaries living in the community in the U.S. due to missingness.
†The 'Other' race category includes other single races not of Hispanic origin, Two or More Races, or Unknown Races. See the Glossary entry for race/ethnicity information.

**Exhibit 6.1.2:** Estimated Number of Female Community Medicare Beneficiaries by Race and Age, in the 2021 MCBS Survey File PUF*

| Race | Age as of 12/31/2021 | Weighted Count |
|---|---|---|
| White non-Hispanic | Under 65 years | 2,294,969 |
| | 65-74 years | 12,666,813 |
| | 75+ years | 9,771,839 |
| Black non-Hispanic | Under 65 years | 726,544 |
| | 65-74 years | 1,667,733 |
| | 75+ years | 1,019,252 |
| Hispanic | Under 65 years | 500,079 |
| | 65-74 years | 1,624,984 |
| | 75+ years | 933,145 |
| Other† | Under 65 years | 287,649 |
| | 65-74 years | 1,095,044 |
| | 75+ years | 709,528 |

SOURCE: 2021 MCBS Survey File PUF, weighted counts.

*Weighted counts may not sum to the total of beneficiaries living in the community in the U.S. due to missingness.
†The 'Other' race category includes other single races not of Hispanic origin, Two or More Races, or Unknown Races. See the Glossary entry for race/ethnicity information.

**Exhibit 6.1.3:** Estimated Number of Community Medicare Beneficiaries by Race and Age, in the 2021 MCBS Survey File PUF*

| Group | Subgroup | Weighted Count |
|---|---|---|
| Overall Total | | 60,921,141 |
| Sex | Male Total | 27,623,562 |
| | Female Total | 33,297,579 |
| Race | White non-Hispanic Total | 45,625,661 |
| | Black non-Hispanic Total | 6,115,187 |
| | Hispanic Total | 5,234,245 |
| | Other Total† | 3,946,049 |

SOURCE: 2021 MCBS Survey File PUF, weighted counts.

*Weighted counts may not sum to the total of beneficiaries living in the community in the U.S. due to missingness.
†The 'Other' race category includes other single races not of Hispanic origin, two or more races, and unknown races. See the Glossary entry for race/ethnicity information.

## 6.2 Targeted Population and Sampling Strata

The targeted population for the MCBS consisted of persons enrolled in one or both parts of the Medicare program, that is, Part A or Part B, as of December 31 of the applicable sample-selection year, and whose address on the Medicare files is in one of the 48 contiguous states (excludes Alaska and Hawaii) or the District of Columbia. For example, for Fall Rounds 2018, 2019, 2020, and 2021 (the four rounds in which the 2018,

2019, 2020, and 2021 Panels included in the 2021 MCBS data were selected), the targeted population included those individuals enrolled as of December 31 of 2018, 2019, 2020, and 2021, respectively.

Additionally, in the 2018, 2019, 2020, and 2021 Panels, beneficiaries residing within the U.S. who were Hispanic (based on a Hispanic ethnicity classification code in the Medicare enrollment data; see Eicheldinger[12] for more details) were oversampled to improve precision of estimates for this group.[13] For more information on the sampling strata, please see Section 6 of the *Survey File Data User's Guide*: https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS/Codebooks.

Exhibit 6.2.1 displays the number of beneficiaries included in the 2021 MCBS Survey File PUF, by age and ethnicity.

**Exhibit 6.2.1:** 2021 Panel of Selected Beneficiaries by U.S. Hispanic and U.S. Non-Hispanic Ethnicity Classification and Age Category*

| Age Category as of 12/31/2021 | TOTAL Sample Size | TOTAL Weighted | Hispanic Sample Size | Hispanic Weighted | Non-Hispanic Sample Size | Non-Hispanic Weighted |
|---|---|---|---|---|---|---|
| Under 65 years | 2,156 | 7,654,400 | 270 | 961,270 | 1,886 | 6,693,130 |
| 65-74 years | 4,508 | 31,699,310 | 507 | 2,715,395 | 4,001 | 28,983,915 |
| 75+ years | 6,119 | 21,567,431 | 619 | 1,557,579 | 5,500 | 20,009,852 |
| Total | 12,783 | 60,921,141 | 1,396 | 5,234,245 | 11,387 | 55,686,896 |

SOURCE: 2021 MCBS Survey File PUF.

*Weighted counts may not sum to the total of beneficiaries living in the community in the U.S. due to missingness.

## 6.3 Primary and Secondary Sampling Units

All of the panels in the 2021 data releases are distributed across the subset of 104 PSUs from the redesigned sample of 107 PSUs selected in 2001. These PSUs are a representative, national sample of beneficiaries who are geographically dispersed throughout metropolitan areas and groups of non-metropolitan counties. Recall that SSUs are census tracts or groups of contiguous tracts within the selected PSUs.

## 6.4 Sample Selection

The MCBS sampling design provides nearly self-weighting (i.e., equal probabilities of selection) samples of beneficiaries within each of the 14 sampling strata. Within the selected PSUs and SSUs, a systematic sampling scheme with random starts is employed for selecting beneficiaries.[14] For each Continuing beneficiary, the survey questions corresponding to the Survey File data release are administered in all three rounds of the collection year. For beneficiaries new to the MCBS, the survey questions are administered as part of the initial fall Baseline interview.

---

[12] Celia Eicheldinger and Arthur Bonito, "More Accurate Racial and Ethnic Codes for Medicare Administrative Data," *Health Care Financing Review* 29, no. 3 (2008): 27-42.

[13] Oversampling of Hispanic beneficiaries has been conducted throughout the MCBS and has evolved over time. See prior *MCBS Methodology Reports* for more information: https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS/Codebooks

[14] The MCBS 2021 Panel was drawn by systematic random sampling with probability proportional to probabilities of selection with an independently selected random start within each PSU. For more information on this sampling method, please see the *MCBS Methodology Report*, available at: https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS/Codebooks.

# 7. TECHNICAL NOTES ON USING THE DATA

## 7.1 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.

Each segment of the MCBS Survey File PUF includes ever enrolled, full sample cross-sectional weights (Fall: PUFFWGT; Winter: PUFWWGT; Summer: PUFSWGT). The MCBS Survey File PUF segments do not include the continuously enrolled cross-sectional weights in order to protect the confidentiality of the beneficiaries. The continuously enrolled cross-sectional weights are available, however, in the LDS.

The ever enrolled cross-sectional weights apply to both the Continuing sample (beneficiaries sampled between 2018-2020) and to the Incoming Panel sample (beneficiaries sampled in 2021). These weights are intended for use in cross-sectional statistics involving the total (combined) Fall 2021 sample. Each weight is greater than zero for all beneficiaries on the file. The ever enrolled cross-sectional weights should be used to make estimates of parameters for the Medicare population who were enrolled at any point in 2021 (i.e., the ever enrolled population).

To generate estimates using the data from merged seasonal segments, the data user must always use the weights that correspond to the segment that is the smaller subset of the other. The samples for the Winter and Summer segments are subsets of the Fall segment sample. Thus, when generating estimates from a merged Summer segment and Fall segment analytic file, the data user must use the Summer segment weights. There are no weights that support joint analysis between the Summer segment and Winter segment, as each segment has a different set of beneficiaries included. To permit the calculation of random errors due to sampling, 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 Survey File PUF can be used to calculate standard errors of the sample-based estimates. The replicate cross-sectional weights in the Fall segment are labeled PUFF001 through PUFF100 corresponding to the ever enrolled weight PUFFWGT. The replicate cross-sectional weights in the Winter segment are labeled PUFW001 through PUFW100 corresponding to the ever enrolled weight PUFWWGT. The replicate cross-sectional weights in the Summer segment are labeled PUFS001 through PUFS100 corresponding to the ever enrolled weight PUFSWGT.

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, the variance estimation method of balanced repeated replication (BRR) using Fay's adjustment of 0.3 is recommended. Sample code in SAS, STATA, and R for estimating statistics can be found in Appendix B. Analysis of subgroups should utilize the domain functions within the statistical package of the data user's choice (e.g., the DOMAIN statement in SAS, or the OVER function in STATA); restricting the sample to the subgroup and then performing an analysis would lead to slightly biased point estimates and estimates of variance.

## 7.2 Item Non-Response

As in any other survey, some respondents could not, or would not, supply answers to some questions.[15] Item non-response rates are generally low in the MCBS data, but the researcher still needs to be aware of the missing data and be cautious about patterns of non-response.[16] The calculation of the study-wide response rates generally follows the guidelines specified in the American Association for Public Opinion Research (AAPOR) and OMB. For the ever enrolled cross-sectional sample represented by the MCBS 2021 Survey File, the calculated overall response rate was 50.8%. 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 the 2021 MCBS Survey File PUF, which excludes beneficiaries for whom only Facility interviews were conducted during the data collection year.

## 7.3 Subgroup Analysis

When analyzing survey data, researchers are often interested in focusing their analyses on specific subgroups of the full population sample (e.g., Medicare beneficiaries aged 65 and over, Hispanics, or females). A common pitfall when performing sub-group analysis of survey data when variance estimation methods such as Taylor-series are used is to delete or exclude observations not relevant to the subgroup of interest. Standard errors for MCBS estimates are most accurate when the analytic file includes all beneficiaries. However, when replicate weights are used for variance estimation, deleting observations not relevant to the subgroup of interest prior to analyzing the subgroup will still produce unbiased standard errors. Almost all statistical packages provide the capability to limit the analysis to a subgroup of the population.

The Taylor Series linearization method of variance estimation is not recommended for subgroup analysis with MCBS data because accidentally excluding any observation in the sample while conducting the subgroup analysis using this variance estimation method will result in biased standard error estimates. Variance estimation using the Taylor Series linearization method for subgroup analyses requires a "domain" or "subgroup" statement (available in most statistical packages) to account for estimated domain sizes (i.e., uncertainty in the denominator). The recommended method of variance estimation for subgroup analysis is the BRR method; which does not require any special subgroup considerations. The BRR method allows the researcher to subset data to a subgroup of interest and still produce unbiased standard error estimates.

## 7.4 Example Research Questions

Exhibit 7.4.1 presents example research questions by topic, differentiating between those that can be addressed by the MCBS Survey File PUF or MCBS Survey File LDS, and those that can only be addressed by the MCBS Survey File LDS. These research questions are intended to illustrate the types of analyses researchers can perform using either the MCBS Survey File PUF or MCBS Survey File LDS, and are not meant to be a comprehensive list of possible research questions that can be answered with these data. The variables that are unique to the MCBS Survey File LDS and not contained in the MCBS Survey File PUF are italicized below.

---

[15] This is different from when an individual refuses to participate in the survey altogether, which is called unit non-response. Unit non-response is discussed in detail in the *MCBS Methodology Report*, Section 9.

[16] In the LDS files, item non-response types are indicated by missing type codes in SAS, including refusal to answer, don't know the answer, and invalid skip. The code .D represents a "don't know" response, the code .R represents a "refused" response, and .N represents an "invalid skip" response.

**Exhibit 7.4.1:** Example Research Questions That Can be Answered Using the MCBS Survey File PUF or MCBS Survey File LDS

| Topic | Example Research Questions Addressed Using the MCBS Survey File PUF or MCBS Survey File LDS | Example Research Questions Addressed Using the MCBS Survey File LDS |
|---|---|---|
| **Quality of Patient Experience** | Are there differences in Medicare beneficiaries' quality of patient experience across socio-demographic characteristics? | Are there differences in Medicare beneficiaries' quality of patient experience between those with *limited English proficiency* and those who are proficient in English? |
| **Access to Care** | Among Medicare beneficiaries, are there differences in access to care by income (below or above $25,000) and level of education? | Among Medicare beneficiaries, are there differences in access to care by *employment status*? |
| **Preventive Care** | Are there differences in receipt of preventive care (e.g., pneumococcal vaccination) by age? | Are there differences in receipt of preventive care (e.g., pneumococcal vaccination) by *patient activation*; that is, the degree to which beneficiaries actively participate in their own health care and decisions concerning that health care? |
| **Cost and Utilization** | Among Medicare beneficiaries, are there differences in the average number of inpatient hospital stays between 2013 and 2021 (MCBS Survey File PUF data were first available in 2013)? | Are there changes in Medicare beneficiaries' *out-of-pocket costs over the last 10 years* (between 2011 and 2021)? |
| **Health Behaviors or Social Determinants of Health** | Are there differences in the percentage of Medicare beneficiaries who smoke cigarettes or consume excessive amounts of alcohol by socio-demographics? | Are there differences in the percentage of Medicare beneficiaries who have *used e-cigarettes* by socio-demographics? |
| **Health Status and Functioning** | Which disease conditions are more common among Medicare beneficiaries who had falls requiring medical help? | Are there differences in the disease conditions associated with falls *between 2011 and 2021*? |
| **Housing Characteristics** | What is the profile (e.g., socio-demographic characteristics and disease conditions) of Medicare beneficiaries with accessibility modifications (e.g., a ramp, bathroom modification and/or railing) in the house? | Are *specific IADL or ADL limitations* associated with accessibility modifications among Medicare beneficiaries? |

# 8. REFERENCES

Eicheldinger, Celia, and Arthur Bonito. "More accurate racial and ethnic codes for Medicare administrative data." *Health care financing review* 29, no. 3 (2008): 27-42.

# APPENDICES

# 9. APPENDICES

## Appendix A: Glossary

**Baseline interview:** The initial questionnaire administered in the fall round of the year the beneficiary is selected into the sample (interview #1).

**Beneficiary:** Beneficiary refers to a person receiving Medicare services who may or not be participating in the MCBS.[17] Beneficiary may also refer to an individual selected from the MCBS sample about whom the MCBS collects information.

**Community component:** Survey administered for beneficiaries living in the community (i.e., not in a long-term care facility such as a nursing home) during the reference period covered by the MCBS interview. An interview may be conducted with the beneficiary or a proxy.

**Continuing interview:** The questionnaire administered as beneficiaries progress through the study (interviews #2-11).

**Continuously enrolled (aka always enrolled):** A Medicare beneficiary who was enrolled in Medicare from the first day of the calendar year until the fall interview and did not die prior to the fall round. This population excludes beneficiaries who dis-enrolled or died prior to their fall interview, residents of foreign countries, and residents of U.S. possessions and territories.

**Core sections:** These sections of the MCBS Questionnaire are of critical purpose and policy relevancy to the MCBS. They may be fielded every round or on a seasonal basis.

**Crossover:** A beneficiary who enters a long-term care facility setting (e.g., nursing homes) or who alternates between a community and a facility setting.

**Ever enrolled:** A Medicare beneficiary who was enrolled at any time during the calendar year including people who dis-enrolled or died prior to their fall interview. Excluded from this population are residents of foreign countries and of U.S. possessions and territories.

**Facility component:** Survey administered for beneficiaries living in facilities, such as long-term care nursing homes or other institutions, during the reference period covered by the MCBS interview. Interviewers conduct the Facility component with staff members located at the facility (i.e., Facility respondents); beneficiaries are not interviewed if they reside at a facility.

**Incoming Panel sample (formerly known as Supplemental Panel):** A statistically sampled group of beneficiaries that enter the MCBS in the fall round of a data collection year. One panel is retired at the conclusion of each winter round, and a new panel is selected to replace it each fall round. Panels are identified by the data collection year (e.g., 2021 Panel) in which they were selected.

**Long-term care facility:** A facility that provides rehabilitative, restorative, and/or ongoing skilled nursing care to patients or residents in need of assistance with activities of daily living.

**Medicare:** Medicare is the federal health insurance program for people who are 65 or over, certain younger people with disabilities, and people with End-Stage Renal Disease (permanent kidney failure requiring dialysis or a transplant, sometimes called ESRD). The different parts of Medicare help cover specific services:

---

[17] https://www.cms.gov/Medicare/Medicare-General-Information/MedicareGenInfo/index.html.

- Hospital Insurance (Part A): covers inpatient hospital stays, care in a skilled nursing facility, hospice care, and some home health care.

- Medical Insurance (Part B): covers certain doctors' services, outpatient care, medical supplies, and preventive services.

- Medicare Advantage (Part C): an alternative to coverage under traditional Medicare (Parts A and B), a health plan option similar to a Health Maintenance Organization (HMO) or Preferred Provider Organization (PPO) administered by private companies.

- Prescription Drug Coverage (Part D): additional, optional coverage for prescription drugs administered by private companies.

For more information, please visit the Medicare.gov website at https://www.medicare.gov/sign-up-change-plans/decide-how-to-get-medicare/whats-medicare/what-is-medicare.html.

**Medicare Advantage (MA):** Medicare Advantage Plans, sometimes called "Part C" or "MA Plans," are offered by private companies under contracts with Medicare. In addition, other managed care plans are offered by private companies under contracts with Medicare under different parts of the Medicare statute. These Medicare managed care plans generally cover Medicare Part A and/or Part B benefits and are paid on either a risk-based capitated basis (MA plans) or on a reasonable cost basis (cost plans and health care prepayment plans).

**Medicare beneficiary:** See Beneficiary.

**Panel:** See Incoming Panel sample.

**Primary Sampling Unit (PSU):** PSU refers to sampling units that are selected in the first (primary) stage of a multi-stage sample ultimately aimed at selecting individual elements (Medicare beneficiaries in the case of MCBS). PSUs are made up of major geographic areas consisting of metropolitan areas or groups of rural counties.

**Proxy:** Beneficiaries who were too ill, or who could not complete the Community interview for other reasons, were asked to designate a proxy, someone very knowledgeable about the beneficiary's health and living habits. In most cases, the proxy was a close relative such as the spouse/partner (if applicable) or a son or daughter. In a few cases, the proxy was a non-relative like a close friend or caregiver. In addition, a proxy was utilized if a beneficiary had been reported as deceased during the current round's reference period or if a beneficiary who was living in the community in the previous round had since entered into a long-term care facility. Proxy interviews are only used for the Community interview, as the Facility interview is conducted with a staff member located at the facility (see definition of "Facility component").

**Race/ethnicity:** Hispanic origin and race are two separate and distinct categories. Persons of Hispanic origin may be of any race or combination of races. Hispanic origin includes persons of Mexican, Puerto Rican, Cuban, Central and South American, or Spanish origin. For the MCBS, responses to beneficiary race and ethnicity questions are reported by the respondent. More than one race may be reported. For conciseness, the text, tables, and figures in this document use shorter versions of the terms for race and Hispanic or Latino origin specified in the Office of Management and Budget 1997 Standards for Data on Race and Ethnicity. Beneficiaries reported as White and not of Hispanic origin were coded as White non-Hispanic; beneficiaries reported as Black/African-American and not of Hispanic origin were coded as Black non-Hispanic; beneficiaries reported as Hispanic, Latino/Latina, or of Spanish origin, regardless of their race, were coded as Hispanic. The "Other" race category includes other single races not of Hispanic origin (including American Indian or Alaska Native, Asian, Native Hawaiian or Other Pacific Islander), Two or More Races, or Unknown Races.

**Respondent:** Respondent refers to a person who answers questions for the MCBS; this person can be the beneficiary, a proxy, or a staff member located at a facility where the beneficiary resides (i.e., the Facility respondent). If the respondent is a proxy, they answer questions about the beneficiary rather than themself.

**Round:** The MCBS data collection period. There are three distinct rounds each year; winter (January through April); summer (May through August); and fall (September through December).

**Secondary Sampling Unit (SSU):** SSUs are made up of census tracts or groups of tracts within the selected PSUs.

**Topical sections:** Sections of the MCBS Questionnaire that collect information on special interest topics. They may be fielded every round or on a seasonal basis. Specific topics may include housing characteristics, drug coverage, and knowledge about Medicare.

**Ultimate Sampling Unit (USU):** USUs are Medicare beneficiaries selected from within the selected SSUs.

# Appendix B: Technical Appendix – Sample Code and Output

Please note that the code examples below use the Survey File PUF Fall segment weights, which begin with the prefix "PUFF." The data user should use the Survey File PUF Fall segment weights only if they are using data from the Survey File PUF Fall segment that have not been merged with data from any other segment. If the data user is analyzing data from the winter, summer, or a combination of Survey File PUF segments, please see the discussion in section 7.1 of this document as to which weights should be used.

## *SAS Analysis Statements*

**Cross-tabulations**
```
proc surveyfreq data=<Analytic dataset> VARMETHOD = brr (fay=.30);
        table <Var name>  / row  chisq lrchisq;
        weight PUFFWGT;
        repweight PUFF001 - PUFF100;
run;
```

**Subgroup Analysis**
```
proc surveyfreq data=<Analytic dataset> VARMETHOD = brr (fay=.30);
        table <Var name> * <Subgroup variable> / row  chisq lrchisq;
        weight PUFFWGT;
        repweight PUFF001 - PUFF100;
run;
```

## *Stata Analysis Statements*

**Declare dataset as survey sample with replicate weights**
```
svyset _n [pweight= PUFFWGT ], brrweight(PUFF001 - PUFF100) fay(.3) vce(brr) singleunit(missing)
```

**For categorical variables, use:**
```
svy brr,  fay(.3) : tabulate <Var name> <Var name>
```

**For subgroup analysis use:**
```
svy brr, subpop(if <Subgroup>) fay(.3) : tabulate <Var name>, over(<Var name>)
```

## *R Analysis Statements*

**Declare MCBS survey design object with replicate weights**
```
mcbs <- svrepdesign(
  weights = ~PUFFWGT,
  repweights = "PUFF[001-100]+",
  type = "Fay",
  rho = 0.3,
  data = <Source dataset>,
  combined.weights = TRUE
  )
```

**For categorical variables, use:**
```
svytable(~<Var name>,  design=mcbs)
```

**For subgroup analysis use:**
```
mcbs_subgrp <- subset(mcbs, <Subgroup criteria>)
svytable(~<Var name>, design=mcbs_subgrp)
```