

2022 | DATA USER'S GUIDE: SURVEY FILE



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

Version Control Log

Date	Version	Revisions
6/21/2024	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* (this document): 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, particularly for researchers 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*: 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 Guides 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.
- **Glossary:** Formerly included as an appendix in MCBS documentation products, this new standalone resource provides the definitions for common key terms used by the MCBS.

TABLE OF CONTENTS

OVERVIEW OF MCBS DOCUMENTATION	ii
ACRONYM LIST	x
1. INTRODUCTION.....	1
1.1 Contents of the Data User's Guide: Survey File.....	2
2. GENERAL GUIDELINES FOR DATA USE.....	4
2.1 Data Access.....	4
2.2 Guidelines for Citation of Data Source	4
3. WHAT'S NEW FOR DATA YEAR 2022?	6
3.1 Sampling	6
3.2 Questionnaires.....	6
3.2.1 General.....	6
3.2.2 Section-Specific Changes.....	6
3.3 Data Collection	10
3.4 Documentation	10
3.5 Data Processing	10
4. SURVEY OVERVIEW	15
4.1 Design of MCBS	15
4.2 Sample Design.....	16
4.3 Case Types.....	17
4.3.1 Incoming and Continuing Cases	17
4.3.2 Community Interviews and Facility Interviews.....	17
4.4 Completed Interviews.....	18
5. QUESTIONNAIRES	19
5.1 Overview.....	19
5.1.1 Items from Validated Scales	19
5.2 Community Questionnaire.....	20
5.2.1 Baseline Interview	21
5.2.2 Continuing Interview.....	22
5.2.3 Core Questionnaire Sections	24
5.2.4 Topical Questionnaire Sections.....	32

5.2.5 Community Questionnaire Section Rotation within a Data Year.....	34
5.3 Facility Instrument	34
5.3.1 Facility Baseline Interview	36
5.3.2 Facility Continuing Interview.....	37
5.3.3 Facility Continuing Core Sections.....	38
5.3.4 Facility Topical Questionnaire Sections.....	41
6. SAMPLING	42
6.1 Medicare Population Covered by the 2022 MCBS Data	42
6.2 Targeted Population and Sampling Strata	46
6.2.1 Eligibility: Medicare Population Covered by the 2022 LDS.....	47
6.3 Three-Stage Cluster Design	47
6.4 Sample Selection	47
7. DATA PRODUCTS & DOCUMENTATION.....	49
7.1 Contents of Data Release	49
7.1.1 2022 MCBS Survey File.....	50
7.1.2 2022 MCBS Cost Supplement File.....	58
7.1.3 Using the Data.....	61
7.1.4 Research Claims Files.....	61
7.2 Which File Do I Need?	61
7.2.1 Survey File Only.....	62
7.2.2 Using Both Survey File and Cost Supplement File	62
7.2.3 Using Both Community and Facility Data.....	62
8. FILE STRUCTURE	64
8.1 LDS Specifications.....	64
8.2 File Structure.....	64
8.2.1 Sort Order for Merging the Survey File LDS Segments	64
8.3 Claims Files	65
8.3.1 Utilization Detail Records.....	65
9. DATA FILE DOCUMENTATION.....	66
9.1 LDS Contents.....	66
9.2 LDS Components	66
9.2.1 Codebooks	66
9.2.2 Questionnaires.....	67
9.2.3 Data User Resources.....	67

9.3 Data Edits and Imputation	67
9.3.1 Data Edits	67
9.3.2 Imputation	68
9.4 Weighting.....	69
9.4.1 Preparing Statistics (Using the Full Sample Weights)	69
9.4.2 Special Topical Segment Weights	71
9.5 Using the Data.....	78
9.5.1 Merging Segments within 2022	78
9.6 Variance Estimation (Using the Replicate Weights)	78
9.6.1 Variables Available for Variance Estimation	78
9.6.2 Variance Estimation for Analyses of Single Year of MCBS	79
9.6.3 Subgroup Analysis.....	79
9.7 Combining Multiple Years of Data.....	79
9.7.1 Longitudinal Analysis.....	79
9.7.2 Repeated Cross-Sectional or Pooled Analysis.....	80
10. DATA FILE NOTES	83
10.1 Global Information	83
10.1.1 BASEID	83
10.1.2 Missing Values	83
10.1.3 Derived and Administrative Variables	83
10.1.4 Initial Interview Variables	83
10.1.5 Ever Variables.....	84
10.1.6 Data Editing.....	84
10.1.7 Other Specify Questions	84
10.1.8 Interview Mode Indicator.....	85
10.1.9 Consistency with Medicare Program Statistics.....	85
10.1.10 Do Administrative Data Override Survey-Reported Data?.....	85
10.2 Survey File Segment Information.....	85
10.3 Survey File Segment Descriptions	85
10.3.1 Access to Care (ACCESSCR).....	85
10.3.2 Access to Care, Medical Appointments (ACCSSMED)	86
10.3.3 Administrative Utilization Summary (ADMNUTLS)	87
10.3.4 Assistance (ASSIST)	93
10.3.5 Chronic Conditions (CHRNCOND)	94
10.3.6 Chronic Conditions Flags (CHRNCDL)	95
10.3.7 Chronic Pain (CHRNPAIN)	96
10.3.8 Cognitive Measures (COGNFUNC).....	96
10.3.9 COVID-19 Experiences (COVIDEXP).....	96

10.3.10 COVID-19 Topical (COVIDTOP)	97
10.3.11 Demographics (DEMO)	98
10.3.12 Diabetes (DIABETES).....	103
10.3.13 Facility Assessments (FACASMNT)	104
10.3.14 Facility Characteristics (FACCHAR).....	108
10.3.15 Falls (FALLS).....	109
10.3.16 Food Insecurity (FOODINS).....	109
10.3.17 General Health (GENHLTH)	109
10.3.18 Health Insurance Summary (HISUMRY)	110
10.3.19 Health Insurance Timeline (HITLINE).....	114
10.3.20 Household Characteristics (HHCHAR).....	116
10.3.21 Income and Assets (INCASSET).....	117
10.3.22 Interview Characteristics (INTERV).....	119
10.3.23 Medicare Advantage Plan Questions (MAPLANQX)	119
10.3.24 Medicare Plan Beneficiary Knowledge (MCREPLNQ)	120
10.3.25 Minimum Data Set (MDS3).....	121
10.3.26 Mental Health (MENTHLTH)	121
10.3.27 Mobility (MOBILITY)	122
10.3.28 Multiple Year Enrollment (MYENROLL)	122
10.3.29 Nagi Disability (NAGIDIS)	122
10.3.30 Nicotine and Alcohol (NICOALCO).....	124
10.3.31 Outcome and Assessment Information (OASIS).....	125
10.3.32 Patient Activation (PNTACT).....	125
10.3.33 Preventive Care (PREVCARE).....	125
10.3.34 Residence Timeline (RESTMLN).....	126
10.3.35 RX Medications (RXMED)	127
10.3.36 Satisfaction with Care (SATWCARE).....	127
10.3.37 Social Determinants of Health (SDOH)	128
10.3.38 Telemedicine (TELEMED)	129
10.3.39 Usual Source of Care (USCARE).....	129
10.3.40 Vision and Hearing (VISHEAR)	130
10.3.41 COVID-19 Facility Beneficiary-Level (FBENCVFL).....	130
10.3.42 COVID-19 Facility Facility-Level (FFACCVFL).....	131
10.3.43 Weights.....	131

11. REFERENCES..... 132

12. APPENDICES 134

Appendix A: MCBS Rounds by Data Year and Season 134

Appendix B: Sample Code 135

Appendix C: Initial Interview Variables 140

Appendix D: Table of Links to MCBS Documentation 142

Appendix E: 2022 MCBS Survey File Segments and Historic RIC Segments..... 143

LIST OF EXHIBITS

Exhibit 3.5.1:	2022 MCBS Content Additions	11
Exhibit 4.1.1:	Typical MCBS Data Collection Year	15
Exhibit 4.1.2:	MCBS Beneficiary Participation Timeline.....	16
Exhibit 4.2:	2018-2022 MCBS Rotating Panel Design.....	17
Exhibit 4.4:	2022 MCBS Fall Round Completed Interviews: Continuing and Incoming Panels	18
Exhibit 5.1:	MCBS Questionnaire Overview	19
Exhibit 5.2:	Overview of the MCBS Community Questionnaire Components	21
Exhibit 5.2.1:	2022 Data Collection Year MCBS Community Questionnaire Flow for Baseline Interview	22
Exhibit 5.2.2:	2022 Data Collection Year MCBS Community Questionnaire Flow for Continuing Interview	23
Exhibit 5.2.3:	2022 Data Year MCBS Community Core Sections by Data File and Data Collection Schedule*	24
Exhibit 5.2.4:	Cost Series Section Overview	27
Exhibit 5.2.5:	Utilization and Cost Section Flow	29
Exhibit 5.2.6:	Example Paths Toward Cost Resolution	29
Exhibit 5.2.7:	Events Collected in the 2022 Data Year	30
Exhibit 5.2.8:	2022 Data Year MCBS Community Topical Sections by Data File and Data Collection Schedule	32
Exhibit 5.2.9:	2022 Data Collection Year MCBS Community Questionnaire Section Rotation	34
Exhibit 5.3:	Overview of the MCBS Facility Instrument.....	35
Exhibit 5.3.1:	2022 Data Collection Year MCBS Facility Instrument Flow for Baseline Interview.....	36
Exhibit 5.3.2:	2022 Data Collection Year MCBS Facility Instrument Flow for Continuing Interviews	37
Exhibit 5.3.3:	2022 Data Year MCBS Facility Core Sections by Data File and Data Collection Schedule*	38
Exhibit 5.3.4:	2022 Data Year MCBS Facility Topical Sections by Data File and Data Collection Schedule.....	41
Exhibit 6.1:	2022 MCBS Composition of Panels Contributing to the LDS Data Files	42

Exhibit 6.1.1:	2022 Total Estimated Number of Medicare Beneficiaries by Sex and Race*	43
Exhibit 6.1.2:	2022 Estimated Number of Male Medicare Beneficiaries by Race and Age*	44
Exhibit 6.1.3:	2022 Estimated Number of Female Medicare Beneficiaries by Race and Age*	45
Exhibit 6.2.1:	2022 MCBS Sampling Strata	46
Exhibit 6.2.2:	2022 Panel of Selected Beneficiaries by Hispanic and Non-Hispanic Ethnicity Classification and Age Category*	47
Exhibit 6.4:	2022 MCBS Sample Selection for the LDS Releases	48
Exhibit 7.1:	2022 Contents of Data Releases.....	50
Exhibit 7.1.1:	2022 MCBS Survey File Segments and Contents	52
Exhibit 7.1.2:	2022 MCBS Cost Supplement File Segments and Contents.....	59
Exhibit 7.2:	MCBS Populations in Data Products	62
Exhibit 8.2.1:	Sort Order by Segment in the Survey File LDS	64
Exhibit 9.3.1:	Data Review and Editing Codes	68
Exhibit 9.4.1:	Crosswalk of 2022 Questionnaire Sections and LDS Segments with Topical Weights.....	71
Exhibit 9.4.2:	2022 MCBS Data Files Summary of Weights	73
Exhibit 9.7.1:	Constructing a Longitudinal Analytic File	80
Exhibit 9.7.2:	Constructing a Repeated Cross-Section or Pooled Analytic File	81
Exhibit 10.3.11.2a:	Primary RUCA (H_RUCA1) Codes, 2010	100
Exhibit 10.3.11.2b:	Secondary RUCA (H_RUCA2) Codes, 2010	101
Exhibit 10.3.13.3a:	FACASMNT Administrative Data Match Outcomes	105
Exhibit 10.3.13.3b:	FACASMNT Variables Populated with Administrative Data	105
Exhibit 10.3.13.3c:	Differences between FACASMNT and MDS3 Data	107
Exhibit 10.3.13.3d:	LDS Segments with Similar Topics for Community and Facility Interviews.....	107
Exhibit 10.3.33.1:	Segment, Questionnaire, and Variable Information for Analyses of 2021-2022 Flu Shot Data.....	126
Exhibit C.1:	Initial Interview Variables.....	140

ACRONYM LIST

ACCESSCR	Access to Care segment
ACCSSMED	Access to Care, Medical Appointments segment
ACQ	Access to Care Questionnaire
ACS	American Community Survey
ADLs	Activities of Daily Living
ADMNUTLS	Administrative Utilization Summary segment
AHRQ	Agency for Healthcare Research and Quality
ASSIST	Assistance segment
ATC	Access to Care
ATSDR	Agency for Toxic Substances and Disease Registry
BQ	Background Questionnaire
BRR	Balanced repeated replication (or Fay's method)
CAPI	Computer-Assisted Personal Interviewing
CASPER	Certification and Survey Provider Enhanced Reports
CCN	CMS Certification Number
CDC	Centers for Disease Control and Prevention
CENWGTS	Continuously enrolled weights
CHRNCDL	Chronic Conditions Flags segment
CHRNCOND	Chronic Conditions segment
CHRNPAIN	Chronic Pain segment
CMS	Centers for Medicare & Medicaid Services
CMQ	Cognitive Measures Questionnaire
COGNFUNC	Cognitive Measures segment
COVIDTOP	COVID-19 Topical segment
COVIDEXP	COVID-19 Experiences segment
CSEVRWGT	Cost Supplement File Ever Enrolled weights
CSL2WGTS	Cost Supplement File Longitudinal weights (2-year)
CSL3WGTS	Cost Supplement File Longitudinal weights (3-year)
CPS	Charge Payment Summary Questionnaire
CV	COVID-19 Beneficiary Supplement Questionnaire
CVQ	COVID-19 Questionnaire
DEMO	Demographics segment
DIABETES	Diabetes segment
DIQ	Demographics and Income Questionnaire
DME	Durable Medical Equipment segment
DUA	Data Use Agreement
DUE	Dental Utilization Events segment
DVH	Dental, Vision, and Hearing Utilization Questionnaire
ENS	Enumeration Summary Questionnaire
EOB	Explanation of Benefit Statements
EPPE	Enterprise Privacy Policy Engine
ERQ	Emergency Room Utilization Questionnaire

ERS	Economic Research Service
ESRD	End-stage renal disease
EVRWGTS	Ever enrolled population weights
EX	Expenditures Questionnaire
FACASMNT	Facility Assessments segment
FACCHAR	Facility Characteristics segment
FAE	Facility Events segment
FALLS	Falls segment
FBENCVFL	COVID-19 Facility Beneficiary-Level segment
FC	COVID-19 Facility-Level Questionnaire
FFACCVFL	COVID-19 Facility Facility-Level segment
FFS	Fee-for-Service
FOODINS	Food Insecurity segment
FQ	Facility Questionnaire
GAD	Generalized Anxiety Disorder screening tool (GAD-2)
GENHLTH	General Health segment
HAQ	Housing Characteristics Questionnaire
HFQ	Health Status and Functioning Questionnaire
HHCHAR	Household Characteristics segment
HHQ	Home Health Utilization Questionnaire
HHS	Home Health Summary Questionnaire
HIPAA	Health Insurance Portability and Accountability Act
HIQ	Health Insurance Questionnaire
HISUMRY	Health Insurance Summary segment
HITLINE	Health Insurance Timeline segment
HMO	Health Maintenance Organization
HS	Health Status
HUE	Hearing Utilization Events segment
IADLs	Instrumental Activities of Daily Living
IAQ	Income and Assets Questionnaire
ID	Identification
IN	Introduction Questionnaire
INCASSET	Income and Assets segment
INQ	Introduction Questionnaire
INTERV	Interview Characteristics segment
IPE	Inpatient Hospital Events segment
IPQ	Inpatient Hospital Utilization Questionnaire
IRB	Institutional Review Board
IRQ	Interviewer Remarks Questionnaire
IUE	Institutional Events segment
IUQ	Institutional Utilization Questionnaire
KNQ	Beneficiary Knowledge and Information Needs Questionnaire
LDS	Limited Data Set(s)
LNG2WGTS	Survey File Longitudinal weights (2-year)
LNG3WGTS	Survey File Longitudinal weights (3-year)

LNG4WGTS	Survey File Longitudinal weights (4-year)
MA	Medicare Advantage
MAPLANQX	Medicare Advantage Plan Questions segment
MBQ	Mobility of Beneficiaries Questionnaire
MCBS	Medicare Current Beneficiary Survey
MCREPLNQ	Medicare Plan Beneficiary Knowledge segment
MDS	Minimum Data Set
MDS3	Minimum Data Set segment
MYENROLL	Multiple Year Enrollment segment
MENTHLTH	Mental Health segment
MMA	Medicare Prescription Drug, Improvement, and Modernization Act of 2003
MOBILITY	Mobility segment
MPE	Medical Provider Events segment
MPQ	Medical Provider Utilization Questionnaire
NAGIDIS	Nagi Disability segment
NICOALCO	Nicotine and Alcohol segment
NORC	NORC at the University of Chicago
NSQ	No Statement Charge Questionnaire
OASIS	Outcome and Assessment Information segment
OEDA	Office of Enterprise Data and Analytics
OMQ	Other Medical Expenses Utilization Questionnaire
OPE	Outpatient Hospital Events segment
OPQ	Outpatient Utilization Questionnaire
PAQ	Patient Activation Questionnaire
PDP	Prescription Drug Plan
PHQ	Patient Health Questionnaire depression screening tool (PHQ-9)
PM	Prescription Medicine
PME	Prescribed Medicine Events segment
PMQ	Prescribed Medicine Questionnaire
PNTACT	Patient Activation segment
PPIC	Patient Perceptions of Integrated Care Questionnaire
PPO	Preferred Provider Organization
PREVCARE	Preventive Care segment
PS	Person Summary segment
PSQ	Post-Statement Charge Questionnaire
PSU	Primary Sampling Units
PUF	Public Use File
PVQ	Preventive Care Questionnaire
PXQ	Physical Measures Questionnaire
RESTMLN	Residence Timeline segment
RH	Residence History
RIC	Record Identification Code
RUCA	Rural-Urban Commuting Area
RXMED	RX Medications segment

RXQ	Drug Coverage Questionnaire
SAS	Statistical Analysis System
SATWCARE	Satisfaction with Care segment
SCF	Sample Control File
SDI	Social Deprivation Index
SDOH	Social Determinants of Health segment
SCQ	Satisfaction with Care Questionnaire
SNF	Skilled Nursing Facility
SS	Service Summary segment
SSU	Secondary Sampling Units
STQ	Statement Cost Series Questionnaire
SVI	Social Vulnerability Index
TELEMED	Telemedicine segment
TLQ	Telemedicine Questionnaire
US	Use of Health Services Questionnaire
USCARE	Usual Source of Care segment
USDA	U.S. Department of Agriculture
USQ	Usual Source of Care Questionnaire
USU	Ultimate Sampling Unit
VISHEAR	Vision and Hearing segment
VRDC	Virtual Research Data Center
VUE	Vision Utilization Events segment

1. INTRODUCTION

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 Medicare Current Beneficiary Survey (MCBS) consists of a representative national sample of the Medicare population sponsored by the Centers for Medicare & Medicaid Services (CMS).¹ 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 on the Medicare program and plays an essential role in monitoring and evaluating 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 64 and below with certain disabling conditions, residing in the United States. Interviews are usually conducted in-person using computer-assisted personal interviewing (CAPI). However, due to the coronavirus disease 2019 (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. The MCBS has conducted continuous data collection 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.

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.

Annually, CMS releases four sets of files – two Microdata Public Use Files (PUFs) and two Limited Data Sets (LDS). The LDS releases are referred to as the Survey File and the Cost Supplement File. The data within the LDS releases are organized into data segments. The Survey File serves as a stand-alone research file and is generally released 18 months after the close of the calendar year for that data collection cohort. Some data for the Survey File are collected into the next calendar year to provide a complete picture of beneficiaries' health and well-being for analysis. For example, income and assets data are collected through the summer into the next calendar year. The Cost Supplement File is usually released approximately three months after the Survey File, when data collection has ended and final administrative and claims data for that calendar year become available. For the 2019 and 2020 data years, a total of three special PUFs were also released that provided data related to COVID-19.

The Survey File contains information on beneficiaries' demographic information, health insurance coverage, self-reported health status and conditions, and responses regarding access to care and satisfaction with care. The Cost Supplement File contains a comprehensive accounting of beneficiaries' health care use, expenditures, and sources of payment. Detailed descriptions of each file, including the contents of the files, file structure,

¹ 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.

information on new variables, key recodes, and administrative sources for select variables are included in each *Data User's Guide* (i.e., Survey File and Cost Supplement File).

Each data release (LDS and PUF) includes a *Data User's Guide* that offers a publicly available, easily searchable resource for data users. Beginning with the 2015 MCBS data release, data user's guides are updated for each new data year to ensure that users have current documentation on the survey design, methods, and estimation as well as MCBS data products. In this Guide, Section 7 ("Data Products and Documentation") provides a crosswalk from historical segments to 2022 segments. Note that for analyses of beneficiaries' health care costs and utilization, data users will need to use the Cost Supplement File in conjunction with the Survey File.

Information on content and access to the MCBS PUFs, including codebooks and additional documentation, can be found at <https://www.cms.gov/data-research/statistics-trends-and-reports/mcbs-public-use-file>

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.² Beneficiary may also refer to an individual selected from the MCBS sample about whom the MCBS collects information.
- Respondent is the 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).

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

1.1 Contents of the Data User's Guide: Survey File

The content of the Survey File 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 Survey File includes data related to Medicare beneficiaries' access to care, health status, and other information regarding beneficiaries' knowledge, attitudes towards, and satisfaction with their health care. The data release also contains demographic data and information on all types of health insurance coverage as well as Fee-for-Service (FFS) claims data, which provide information on medical services and payments made by Medicare under this plan type.

This Guide contains detailed information about the Survey File and specific background information to help data users understand and analyze the data. A companion *Data User's Guide* focuses on the Cost Supplement File LDS release.³ Data users can access this Guide along with other data documentation at <https://www.cms.gov/data-research/research/medicare-current-beneficiary-survey/data-documentation-codebooks>.

Here is an overview of the contents of the Data User's Guide: Survey File:

- Section 2: General Guidelines for Data Use – This section describes the main requirements for data use.
- Section 3: What's New? – This section describes the key MCBS Questionnaire changes and other highlights and enhancements for the data year.

² <https://www.cms.gov/Medicare/Medicare-General-Information/MedicareGenInfo/index.html>

³ The Cost Supplement LDS and companion *Data User's Guide* is released three to four months after the Survey File LDS.

- Sections 4-9: Overview of the MCBS – These sections provide an overview of the MBCS, including the questionnaires and the file structure. They include a technical description of the specifications and structure of the file and a brief description of the record types in this file.
- Section 10: Data File Notes – This section provides an overview of each file included in the release, a description of derived variables, and any changes from previous releases or special highlights for data users.
- Sections 11-12: References and Appendices – This section provides references and key supporting documentation, including sample programs for data users.

2. GENERAL GUIDELINES FOR DATA USE

The LDS files contain beneficiary-level health information, but exclude specific direct identifiers as outlined in the Health Insurance Portability and Accountability Act of 1996 (HIPAA). LDS files are considered identifiable, even without the inclusion of specific direct identifiers, due to the potential capability to link other sources of data, creating an increased risk of re-identification of individuals. Since the information provided on an LDS is considered identifiable, it also remains subject to the provisions of the Privacy Act of 1974.

2.1 Data Access

All requested LDS files require a signed LDS Data Use Agreement (DUA) between CMS and the data requestor to ensure that the data remain protected against unauthorized disclosure. LDS requestors must show that their proposed use of the data meets the disclosure provisions for research. The research purpose must relate to projects that could ultimately improve the care provided to Medicare patients and policies that govern the care. This type of research includes projects related to improving the quality of life for Medicare beneficiaries, improving the administration of the Medicare program, cost and payment related projects, and the creation of analytical reports. In addition, these research projects must contribute to generalizable knowledge.

Data users can submit an LDS request via a CMS DUA tracking system, the Enterprise Privacy Policy Engine or EPPE. EPPE can be used to initiate a new LDS DUA request or to amend/update an existing LDS DUA.

Questions about LDS files or the process for requesting LDS files can be sent to datauseagreement@cms.hhs.gov. For additional information on data access and the DUA process, including instructions for accessing and using EPPE to make a request, data users can visit the CMS LDS website at <https://www.cms.gov/data-research/files-for-order/data-disclosures-and-data-use-agreements-duas/limited-data-set-lds>.

Administrative processing fees for obtaining the LDS files are \$300 for the 2022 Survey File alone, and \$600 for the 2022 Survey File with the 2022 Cost Supplement File (the Cost Supplement File cannot be acquired separately). The processing of the DUA takes approximately six to eight weeks. Upon approval and payment, CMS releases the data within ten business days, depending on the size of the data request. Data users will receive the data on DVD or flash drive, or via the CMS Virtual Research Data Center (VRDC) for use with SAS® or other statistical software packages; each data release contains multiple files that are linkable through a key identification variable (BASEID).

Questionnaires, codebooks, and Bibliographies for each survey year are available for download on the CMS MCBS website at <https://www.cms.gov/data-research/research/medicare-current-beneficiary-survey>. A link to this documentation is also visible when approved data users log in to the VRDC.

2.2 Guidelines for Citation of Data Source

This communication was printed, published, or produced and disseminated at U.S. taxpayer expense. All material appearing in this document is in the public domain and may be reproduced or copied without permission; citation as to source, however, is appreciated. Accordingly, CMS requests that data users cite CMS and the Medicare Current Beneficiary Survey as the data source in any publications or research based upon these data. Suggested citation formats are below.

Tables and Graphs: The suggested citation to appear at the bottom of all tables and graphs should read:

SOURCE: Centers for Medicare & Medicaid Services, Medicare Current Beneficiary Survey, [Data Product], [Year].

Bibliography: The suggested citation for the *2022 MCBS Data User's Guide* should read:

SOURCE: Centers for Medicare & Medicaid Services. *2022 Medicare Current Beneficiary Survey Data User's Guide: Survey File*. Retrieved from [ADD URL], [YEAR].

Survey Data: The suggested citation for the MCBS survey data files and other documentation should read:

SOURCE: Centers for Medicare & Medicaid Services. Medicare Current Beneficiary Survey, Survey File data. Baltimore, MD: U.S. Department of Health and Human Services, 2022.

SOURCE: Centers for Medicare & Medicaid Services. Medicare Current Beneficiary Survey, Cost Supplement File data. Baltimore, MD: U.S. Department of Health and Human Services, 2022.

3. WHAT'S NEW FOR DATA YEAR 2022?

Below are the highlights and updates for the 2022 data year.

3.1 Sampling

There were no changes to sampling for the 2022 data year.

3.2 Questionnaires⁴

Questionnaire content changes: There were a number of questionnaire sections that were revised in 2022. Note that variable names referenced below are the questionnaire variable names. Data users can view the questionnaire for each data year along with the questionnaire variable names referenced below and question text on the MCBS website at <https://www.cms.gov/data-research/research/medicare-current-beneficiary-survey/questionnaires>.

3.2.1 General

Community Questionnaire

The MCBS introduced several Community Questionnaire updates in 2022 to enhance survey content and data quality, improve interviewer and respondent experience, and address the evolving COVID-19 pandemic.

Additional details about questionnaire content and section-specific changes made in 2022 can be found in Section 3.2.2.

Facility Instrument

The MCBS introduced several Facility Instrument updates in 2022 to streamline content, improve interviewer and respondent experience, and address the evolving COVID-19 pandemic.

Additional details about questionnaire content and section-specific changes made in 2022 can be found in Section 3.2.2.

3.2.2 Section-Specific Changes

Community Questionnaire

Changes implemented for the 2022 Community Questionnaire include the addition of a new questionnaire section, new items, and updates to question text, response options, and respondent universes.⁵

⁴ Variable names referenced in section 3.2 are questionnaire variable names. The names and question text can be viewed on the MCBS website at <https://www.cms.gov/data-research/research/medicare-current-beneficiary-survey/questionnaires>.

⁵ Questionnaire sections CVQ, KNQ, PVQ, TLQ and USQ are administered following the year of interest. Therefore, data collected for these sections in 2023 are released in the 2022 LDS. For more information on the differences between survey administration year versus data year, please refer to Tables 5.2.3 and 5.2.8.

Summary of Item- and Section-Level Questionnaire Revisions

COVID-19 (CVQ)

Several changes were made to the COVID-19 Questionnaire (CVQ):

- Seven items were removed from the CVQ section in Fall 2022 due to low variation in responses within a round, low variability across rounds, and decreased relevancy at the current stage of the pandemic. The series on antibody testing (CVTSTPAY, ANTBDTST, ANTRESLT, ANTWAIT, ANTPAY) and two items on COVID-19 prevention (PREVMASK and PREVGRP) were deleted.
- Also in Fall 2022, question text at SWABRSLT, which collects whether a COVID-19 test found the beneficiary had COVID-19, was modified to instruct beneficiaries to answer in the affirmative if any COVID-19 tests taken during the reference period were positive. To reflect the corresponding change in the universe of respondents at this item, the variable was renamed COVRSLT.
- To align with the 2022 National Health Interview Survey (NHIS),⁶ the CVQ items that collect persistent symptoms of COVID-19 (SMPTFATG, SMPTHEAD, SMPTHRT, SMPTACHE, SMPTCOGH, SMPTDIZZ, SMPTANX, SMPTOTH) were deleted and replaced with the NHIS item collecting symptoms of long COVID-19 (LONGCVD).
- Three items were modified in Fall 2022 to account for the varying reference periods between Baseline and Continuing cases. Each of the three items was split into two new variables to reflect an "ever" reference period for Baseline cases and a "since the date of the last interview" for Continuing cases. Item SUSPECT was replaced with EVRSUS (Baseline) and COVSUS (Continuing), item COVIDEV was replaced with EVRCVTLD (Baseline) and COVTOLD (Continuing), and item COVSWAB was replaced with EVRCVTST (Baseline) and COVTEST (Continuing).
- In Winter 2023, the question text and code list were updated at VACNME, which collects the name of the COVID-19 vaccine the beneficiary received. This item was revised to include the vaccine Novavax after the Food and Drug Administration (FDA) granted its emergency use authorization in July 2022.⁷

Health Insurance (HIQ)

In Winter 2023, several updates were made to improve the Health Insurance Questionnaire (HIQ). The purpose of these changes was threefold: align collection of health insurance information across different plan types, reduce respondent burden by discontinuing collection of detailed information with little analytic utility, and improve the quality of information collected. The full scope of these changes will be reflected in the 2023 Survey File and detailed in the 2023 LDS documentation. However, some information about health insurance plans collected in Winter 2023 with coverage period extending back to 2022 is included in the 2022 Survey File.

Health Status & Functioning (HFQ)

There were four updates to the Health Status & Functioning (HFQ) section in Fall 2022:

- Item OCCHOLES, which collects whether a doctor or other health professional has ever told the beneficiary they have high cholesterol, was previously erroneously administered to Baseline and Continuing cases. In Fall 2022, the universe of respondents for this item was corrected to only Baseline cases and the item was renamed OCCLSTRL. Item YRCHOLES was similarly administered to Baseline and Continuing cases when it was intended to only collect whether Continuing cases were told in the prior year if they had high cholesterol. The administration of this item was corrected to the intended universe and renamed (YRCLSTRL).

⁶ Centers for Disease Control and Prevention, National Center for Health Statistics. (2022). 2022 National Health Interview Survey (NHIS) Questionnaire. https://ftp.cdc.gov/pub/Health_Statistics/NCHS/Survey_Questionnaires/NHIS/2022/EnglishQuest-508.pdf

⁷ <https://ir.novavax.com/press-releases/2022-07-13-U-S-FDA-Grants-Emergency-Use-Authorization-for-Novavax-COVID-19-Vaccine%2C-Adjuvanted-for-Individuals-Aged-18-and-Over>

- The code list at VITSUPYR, which collects what vitamins and dietary supplements the beneficiary has taken during the reference period, was updated to accommodate beneficiaries who only take vitamins and dietary supplements as part of a multivitamin.
- Items TESTDAY, TESTWEEK, TESTMNTN, and TESTYEAR collect how often a beneficiary who has reported diabetes tests their blood for sugar or glucose. In Fall 2022, these items were collapsed into one new item CONTGLUC to accommodate beneficiaries who use a continuous glucose monitoring system to check their blood sugar. The testing frequency items are released in the Survey File as variables D_TSTFRQ and TESTUNIT, and the continuous glucose monitoring item is released as variable CONTGLUC.
- The code list at HYPEDRNK, which asks whether the beneficiary has cut down on drinking alcoholic beverages because of their high blood pressure, was updated to include "Not applicable; respondent does not drink alcohol" to accommodate beneficiaries who do not drink alcoholic beverages.

Beneficiary Knowledge and Information Needs (KNQ)

Four items were added to the Beneficiary Knowledge and Information Needs Questionnaire (KNQ) in Winter 2023:

- Items USEMSP and APPLYMSP were added to gather information about beneficiary participation in the Medicare Savings Program (MSP). The first item (USEMSP) provides a definition of the MSP program and asks if the beneficiary receives assistance from the program. For respondents who say no, a second item (APPLYMSP) asks if the beneficiary applied to their state Medicare office for help with medical expenses. These items are formatted in a way similar to existing MCBS Drug Coverage Questionnaire (RXQ) items about the Low-Income Subsidy (LIS) program.
- Item RGHTAPL was added to understand beneficiary knowledge about the Medicare appeal process and asks the beneficiary if they would know how to file a complaint or an appeal with Medicare if they had concerns about the quality of care they received. This item was drafted by the Beneficiary and Family Centered Care Quality Improvement Organization Program (BFCC-QIO).
- INTERNET, which quantifies beneficiary access to the internet, was added back into KNQ in Winter 2023. This item had previously been fielded as part of the MCBS COVID-19 Supplements. The addition of INTERNET changes the universe of respondents for the subsequent existing questions. Therefore, the variable names of KNETPERS, KNETFRND, and KNETOFTN were changed to USENET, SOMELNET, and OFTNNET, respectively. Additionally, the text fills at KVSTITE, KCOMINTE, KCOMPRES, KCOMAPPO, and KCOMCOMM were updated to reflect the variable name changes.

Prescribed Medicine Utilization (PMQ)

- In Winter 2022, the reference period was corrected from "CURRENT YEAR" to "CURRENT YEAR-1" at SCPMMAIN, which captures the reason for forgone care of prescribed medicines, in the Prescribed Medicine Utilization (PMQ) section.

Drug Coverage (RXQ)

- In Summer 2023, updates were made to the Drug Coverage Questionnaire (RXQ) to accommodate corresponding updates made in the Health Insurance Questionnaire (HIQ) in Winter 2023. The new HIQ structure altered how prescription drug coverage was collected for certain plan types, which subsequently affected the universe of cases eligible for RXQ in Summer 2023.

Usual Source of Care (USQ)

- During Winter 2022 data collection, field interviewers reported that PRVNOMED, which asks if there was one provider who knew about all the medicines the beneficiary was taking, does not apply to beneficiaries who did not take any medications within the past year. To clarify the intent of this item and assist with

administration, interviewer on-screen help text was added to PRVNOMED in Winter 2023 instructing interviewers to probe whether the beneficiary's provider knew they were not taking any medicines.

Facility Instrument

There were several changes made to the Facility Instrument in 2022, including streamlining of the instrument and updates to question text and programming logic.

Background Questionnaire (BQ)

- In the BQ section, BQ15-BHWLIVES collects where the beneficiary's spouse lives and had a text fill that displayed "husband" or "wife" based on the reported sex of the beneficiary. In Winter 2022, this text fill option was replaced with "spouse" at BQ15-BHWLIVES to remove the assumption of the spouse's sex.

COVID-19 Beneficiary (CV)

- Several changes were made within the COVID-19 Beneficiary (CV) section in 2022:
 - ▶ With the approval of COVID-19 vaccine booster doses in the Fall of 2021, it was necessary to build a questionnaire structure that had more flexibility to capture additional vaccine doses and dose details. Therefore, in Winter 2022, a roster-based looping structure was programmed to capture COVID-19 vaccination doses and all associated details, including the date of the dose, the vaccine manufacturer, and the location where the vaccine was administered. In light of this new structure, several new variables were added, and several existing variables were deleted.
 - ▶ Due to low variation in responses within a round and low variability in responses across rounds, the series on antibody testing (ANTICVD and ANTIRE) was deleted in Fall 2022. To accommodate this deletion, text at the CV section introduction screen (CV1-CVDINTRO) was updated to reflect that antibody tests are no longer collected.
 - ▶ Also in Fall 2022, question text at CV2A-TESTRES, which collects whether a COVID-19 test found the beneficiary had COVID-19, was modified to instruct Facility respondents to answer in the affirmative if the beneficiary had any positive COVID-19 tests taken during the reference period. To reflect the corresponding change in the universe of respondents at this item, the variable was renamed CV2B-COVRSLT. To accommodate this update, on-screen interviewer help text was removed at CV2-CVDTEST.

COVID-19 Facility-Level (FC)

- The following changes were made within the COVID-19 Facility-Level (FC) section in 2022:
 - ▶ In Winter 2022, the phrase "coronavirus vaccination" was updated to "COVID-19 vaccine" throughout FC.
 - ▶ Due to low variation in responses within a round and low variability in responses across rounds, the series on facility health care personnel adherence to COVID-19 prevention strategies was deleted in Fall 2022 (FC21-HCPHH, HCPPPE, and HCPDES). To accommodate this update, on-screen text at FC17-ACTINTRO was updated to describe the next series of FC items.

Facility Questionnaire (FQ)

- A few changes were made within the Facility Questionnaire (FQ) section in 2022:
 - ▶ In the Facility Questionnaire (FQ) section, FR3-HIGHRATE and FR3-HIGHPER collect the highest rate and unit for a resident's basic care at the facility, while FR4-LOWRATE and FR4-LOWPER collect the lowest rate and unit for a resident's basic care at the facility. FR3-HIGHRATE and FR3-HIGHPER are administered before FR4-LOWRATE and FR4-LOWPER. In Winter 2022, the rate and unit entered at FR3-HIGHRATE and FR3-HIGHPER were added as a text header at FR4-LOWRATE and FR4-LOWPER, so that the interviewer can be reminded of the rate and unit entered for the highest rate and to help prevent lowest billing rates from being higher than the highest billing rates.

- ▶ Interviewer on-screen help text was added at items collecting total number of long-term care beds in the facility (FA12-TOTLBEDA and FB18-TOTELBED), to provide a definition of long-term care and better assist interviewers in administering these items.

Health Status (HS)

- In Fall 2022, a new variable (HA33A-HA33ACAN) was added in the Health Status (HS) section to ask if cancer was an active diagnosis for interviews conducted at Medicare- or Medicaid-certified facilities. To accommodate this change, question text at HA10B-HA10BCOD was updated to reflect that there may be further items in the HS section that require the MDS.

Residence History (RH)

- The on-screen help text at RHSEX, which collects the beneficiary's sex, previously instructed interviewers to only "ask if not obvious." Starting in Fall 2022, the help text was removed to ensure that the item is read aloud to all Facility respondents.

Use of Health Services (US)

- In Fall 2022, the "SERVICE SUSPENDED DUE TO COVID-19" response option was removed from several items which ask about health care services the beneficiary may have received (OUTMDVST, INMDVST, DENTVST, MENTLVST, PHYSTHPY, PODRTHPY, EDHBSERV, and OTHCPROV). The response option was removed due to decreased relevancy at the current stage of the pandemic, as it is no longer as common for facilities to temporarily or indefinitely suspended certain health care services to prevent the spread of the COVID-19 virus.

3.3 Data Collection

MCBS data collection includes both in-person and phone outreach and interviewing. Mode is determined by efficiency in outreach, beneficiary preference, and local interviewer availability, among other factors. Data quality is monitored to ensure high quality data are collected, regardless of mode.

3.4 Documentation

This 2022 *Data User's Guide* was enhanced with the following content:

- Formerly included as Appendix A of this document, the definitions of common key terms used by the MCBS have been moved into a standalone Glossary document available on the CMS website:
<https://www.cms.gov/data-research/research/medicare-current-beneficiary-survey/data-documentation-codebooks>.

3.5 Data Processing

New and revised content:

For the 2022 Survey File LDS, the MCBS created the following new segments:

- COVIDTOP contains information on COVID-19 vaccination, testing, diagnosis, symptoms, and prevention from Winter 2023 and Summer 2023.
- SDOH contains information on beneficiaries' residential history and social determinants of health.

The 2022 questionnaire changes, as well as updates to administrative or externally sourced data, resulted in the following variables added to the annual release.

Exhibit 3.5.1: 2022 MCBS Content Additions

Location	Questionnaire Section	Variable	Description
CHRNCDL	N/A (admin data)	HF	CLAIM FOR HEART FAILURE THIS BENEFIT YR
CHRNCDL	N/A (admin data)	HFE	1ST DATE EVR CLAIM FOR HEART FAILURE
COVIDEXP	CVQ	CVDOSE	NUMBER OF COVID-19 VACCINE DOSES (EVER)
COVIDEXP	CVQ	VACDATMM	MONTH OF LAST COVID-19 VACCINE DOSE
COVIDEXP	CVQ	VACDATYY	YEAR OF LAST COVID-19 VACCINE DOSE
COVIDEXP	CVQ	VACSITE	SITE OF LAST COVID-19 VACCINE DOSE
COVIDEXP	CVQ	VACNME	BRAND OF LAST COVID-19 VACCINE DOSE
COVIDEXP	CVQ	EVRCVTLD	SP EVER BEEN TOLD THEY LIKELY HAVE CORONAVIRUS
COVIDEXP	CVQ	EVRCVTST	HAS SP EVER BEEN TESTED FOR CORONAVIRUS TO SEE IF INFECTED AT TIME OF TEST
COVIDEXP	CVQ	EVRSUS	SP EVER SUSPECTED THEY HAD CORONAVIRUS
COVIDEXP	CVQ	COVRSLT	RESULT OF ACTIVE COVID-19 TEST
COVIDEXP	CVQ	LONGCVD	SP EXPERIENCED SYMPTOMS 3 MONTHS OR LONGER NOT PRESENT BEFORE HAVING COVID-19
COVIDEXP	CVQ	PRSUMVAC	SP PLANS TO GET COVID-19 VACCINE
COVIDEXP	CVQ	NOEFFECT	REASON FOR NO COVID-19 VACCINE: SIDE EFFECTS
COVIDEXP	CVQ	NOALLERG	REASON FOR NO COVID-19 VACCINE: ALLERGY CONCERN
COVIDEXP	CVQ	NOPROTEC	REASON FOR NO COVID-19 VACCINE: UNSRURE OF COVID-19 VACCINE BENEFIT
COVIDEXP	CVQ	NOBELIEV	REASON FOR NO COVID-19 VACCINE: NOT NEEDED
COVIDEXP	CVQ	NOHADCVD	REASON FOR NO COVID-19 VACCINE: ALREADY HAD COVID-19
COVIDEXP	CVQ	NOHIRISK	REASON FOR NO COVID-19 VACCINE: NOT AROUND HIGH-RISK PEOPLE
COVIDEXP	CVQ	NOMASK	REASON FOR NO COVID-19 VACCINE: MASKS/PRECAUTION INSTEAD
COVIDEXP	CVQ	NOBENE	REASON FOR NO COVID-19 VACCINE: VACCINES DO NOT BENEFIT
COVIDEXP	CVQ	NOSTRONG	REASON FOR NO COVID-19 VACCINE: IMMUNE STRONG ENOUGH
COVIDEXP	CVQ	NODRREC	REASON FOR NO COVID-19 VACCINE: DR. DIDN'T RECOMMEND VACCINE
COVIDEXP	CVQ	NOWAIT	REASON FOR NO COVID-19 VACCINE: WAIT TO SEE IF VACCINES IS SAFE
COVIDEXP	CVQ	NOCOST	REASON FOR NO COVID-19 VACCINE: COST
COVIDEXP	CVQ	NOTSTVAC	REASON FOR NO COVID-19 VACCINE: DO NOT TRUST COVID-19 VACCINE
COVIDEXP	CVQ	NOTHREAT	REASON FOR NO COVID-19 VACCINE: COVID-19 IS NOT A THREAT
COVIDEXP	CVQ	NOHARD	REASON FOR NO COVID-19 VACCINE: DIFFICULT TO GET
COVIDEXP	CVQ	NOFAMFRD	REASON FOR NO COVID-19 VACCINE: FRIENDS/FAMILY ALSO DID NOT

Location	Questionnaire Section	Variable	Description
COVIDEXP	CVQ	NONEEDLE	REASON FOR NO COVID-19 VACCINE: AFRAID OF NEEDLES
COVIDEXP	CVQ	NOBRAND	REASON FOR NO COVID-19 VACCINE: CANNOT GET PREFERRED BRAND
COVIDEXP	CVQ	NOSCHED	REASON FOR NO COVID-19 VACCINE: APPOINTMENT SCHEDULED
COVIDEXP	CVQ	NOCOND	REASON FOR NO COVID-19 VACCINE: HEALTH CONDITION
COVIDEXP	CVQ	NOELIG	REASON FOR NO COVID-19 VACCINE: NOT ELIGIBLE YET
COVIDEXP	CVQ	NOOTH	REASON FOR NO COVID-19 VACCINE: OTHER
COVIDTOP	CVQ	D_CVDOSE	NUMBER OF COVID-19 VACCINES DOSES (EVER); INCLUDES DOSES RECORDED BY SUMMER 2023
COVIDTOP	CVQ	D_RCNTVX	COVID-19 VACCINE DOSE RECORDED SINCE FALL 2022
COVIDTOP	CVQ	D_VACDTM	MONTH OF COVID-19 VACCINE DOSE SINCE FALL 2022
COVIDTOP	CVQ	D_VACDTY	YEAR OF COVID-19 VACCINE DOES SINCE FALL 2022
COVIDTOP	CVQ	D_VACSITE	SITE OF COVID-19 VACCINE DOSE SINCE FALL 2022
COVIDTOP	CVQ	D_VACNME	BRAND OF COVID-19 VACCINE DOSE SINCE FALL 2022
COVIDTOP	CVQ	D_COVSUS	SP SUSPECTED THEY HAD COVID-19 SINCE FALL 2022
COVIDTOP	CVQ	D_COVRSL	RESULTS OF ACTIVE COVID-19 TEST
COVIDTOP	CVQ	D_COVTST	HAS SP BEEN TESTED FOR COVID-19 TO SEE IF INFECTED AT TIME OF TEST SINCE FALL 2022
COVIDTOP	CVQ	D_COVTLD	SP BEEN TOLD THEY LIKELY HAVE COVID-19 SINCE FALL 2022
COVIDTOP	CVQ	D_LONGCV	SP EXPERIENCED SYMPTOMS 3 MONTHS OR LONGER NOT PRESENT BEFORE HAVING COVID-19
COVIDTOP	CVQ	D_PRSMVC	SP PLANS TO GET COVID-19 VACCINE
COVIDTOP	CVQ	D_NOALLE	REASON FOR NO COVID-19 VACCINE: ALLERGY CONCERN
COVIDTOP	CVQ	D_NOHIRI	REASON FOR NO COVID-19 VACCINE: NOT AROUND HIGH-RISK PEOPLE
COVIDTOP	CVQ	D_NOMASK	REASON FOR NO COVID-19 VACCINE: MASKS/PRECAUTION INSTEAD
COVIDTOP	CVQ	D_NOBENE	REASON FOR NO COVID-19 VACCINE: VACCINES DO NOT BENEFIT
COVIDTOP	CVQ	D_NOSTRO	REASON FOR NO COVID-19 VACCINE: IMMUNE STRONG ENOUGH
COVIDTOP	CVQ	D_NOWAIT	REASON FOR NO COVID-19 VACCINE: WAIT TO SEE IF COVID-19 VACCINE IS SAFE
COVIDTOP	CVQ	D_NOHARD	REASON FOR NO COVID-19 VACCINE: DIFFICULT TO GET
COVIDTOP	CVQ	D_NOFAMF	REASON FOR NO COVID-19 VACCINE: FRIENDS/FAMILY ALSO DID NOT
COVIDTOP	CVQ	D_NOBRAND	REASON FOR NO COVID-19 VACCINE: CANNOT GET PREFERRED BRAND
COVIDTOP	CVQ	D_NOCOND	REASON FOR NO COVID-19 VACCINE: HEALTH CONDITION

Location	Questionnaire Section	Variable	Description
COVIDTOP	CVQ	D_NOELIG	REASON FOR NO COVID-19 VACCINE: NOT ELIGIBLE YET
DIABETES	HFQ	CONTGLUC	SP USES CONTINUOUS GLUCOSE MONITOR
FBENCVFL	CV	FONEDOSE	EVER RECEIVED 1 OR MORE COVID-19 VACCINE DOSES
FBENCVFL	CV	FCVDOSE	NUMBER OF COVID-19 VACCINE DOSES (EVER)
FBENCVFL	CV	FVACDATM	MONTH OF LAST COVID-19 VACCINE DOSE
FBENCVFL	CV	FVACDATY	YEAR OF LAST COVID-19 VACCINE DOSE
FBENCVFL	CV	FVACSITE	SITE OF LAST COVID-19 VACCINE DOSE
FBENCVFL	CV	FVACNME	NAME OF LAST COVID-19 VACCINE DOSE
HITLINE	HIQ	D_PREMMON	MONTHLY COST OF PRIVATE HEALTH INSURANCE PLAN PREMIUM
HITLINE	N/A (admin data)	D_PREMN_I	IMPUTATION FLAG FOR D_PREMMON
MAPLANQX	N/A (admin data)	ANHMO_I	IMPUTATION FLAG FOR D_ANHMO PREMIUM VARIABLE
MCREPLNQ	KNQ	APPLYMSP	SP APPLIED TO STATE MEDICARE OFFICE FOR HELP PAYING COSTS
MCREPLNQ	KNQ	INTERNET	SP HAS ACCESS TO INTERNET
MCREPLNQ	KNQ	OFTNNET	HOW OFTEN ACCESS INTERNET FOR INFO?
MCREPLNQ	KNQ	RGHTAPL	SP KNOWS HOW TO FILE A COMPLAINT ABOUT QUALITY OF CARE
MCREPLNQ	KNQ	SOMELNET	SOMEONE ELSE GETS INFO ON INTERNET FOR SP
MCREPLNQ	KNQ	USEMSP	SP USES MEDICARE SAVINGS PLAN TO HELP PAY HEALTH CARE COSTS
MCREPLNQ	KNQ	USENET	SP PERSONALLY USES INTERNET FOR INFORMATION
SDOH	N/A (admin data)	STATE	STATE OF RESIDENCE
SDOH	N/A (admin data)	COUNTY	FIPS COUNTY CODE OF RESIDENCE
SDOH	N/A (admin data)	ZIP	POSTAL ZIP CODE OF RESIDENCE
SDOH	N/A (admin data)	ADDRCHNG	CHANGE OF ADDRESS SINCE PREVIOUS FALL ROUND
SDOH	N/A (admin data)	DISTANCE	MILES BETWEEN LAST YR AND CURRENT YR ADDRESS
SDOH	N/A (admin data)	NUMADDR	TOTAL NUMBER OF UNIQUE ADDRESSES
SDOH	N/A (external) ¹	SDI	SDI BASED ON COUNTY PERCENTILE RANK
SDOH	N/A (external) ²	SVI	SVI BASED ON COUNTY PERCENTILE RANK
SDOH	N/A (external) ³	HOSPBEDS	HOSPITAL BEDS IN COUNTY PER 1,000 POPULATION
SDOH	N/A (external) ³	LTCBEDS	LONG TERM CARE BEDS IN COUNTY PER 1,000 POPULATION
SDOH	N/A (external) ³	MDRATE	TOTAL NUMBER OF NON-FEDERAL MEDICAL DOCTORS IN COUNTY PER 1,000 POPULATION

¹ Sourced from the Social Deprivation Index.

² Sourced from the Centers for Disease Control and Prevention (CDC)/Agency for Toxic Substances and Disease Registry (ATSDR) Social Vulnerability Index.

³ Sourced from the Agency of Healthcare Research and Quality (AHRQ)'s Social Determinants of Health Database.

Weighting:

There were no changes to the weighting processes or deliveries for 2022.

Imputation:

The 2022 Income and Assets imputation added separate variables for monthly earnings from work for the beneficiary and for their spouse or partner. Previously, these variables were combined into a single variable and this variable was imputed on its own.

Health insurance premiums for private plans have been standardized to reflect monthly premiums and, when missing, premium information has been imputed, resulting in the addition of two new variables on HITLINE containing the premium (D_PREMMON) and an imputation flag (D_PREMN_I).

Health insurance premiums have also been imputed for MA plans, resulting in the addition of a new variable on MAPLANQX containing an imputation flag (ANHMO_I) for the premium.

See the *2022 Cost Supplement File Data User's Guide* for information on imputation pertaining to the 2022 Cost Supplement File LDS data.

4. SURVEY OVERVIEW

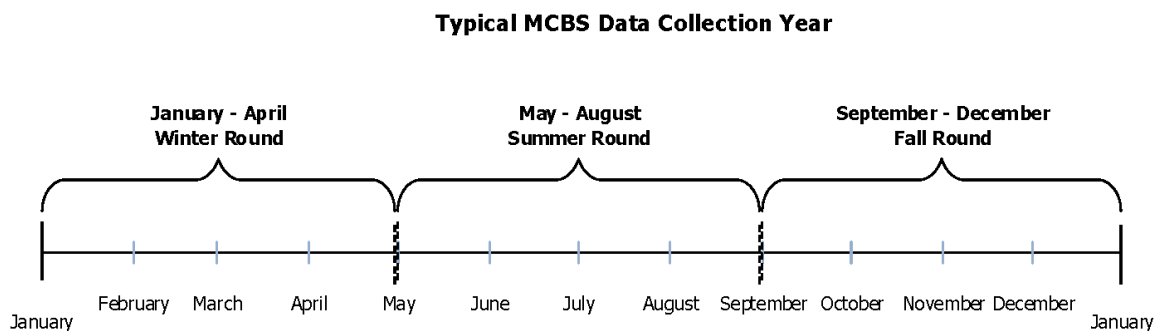
4.1 Design of MCBS

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, this was later determined to be impractical, and beginning in 1994, participation of beneficiaries in the MCBS was limited to no more than four years.

Although limited to a four-year period, MCBS data collection is continual 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 by round configuration (rather than interviewing on an annual basis) is to have shorter periods of recall during the year in order to capture more complete health care costs and utilization from beneficiaries.

The 2022 MCBS data releases reflect data collected from January 2022 through December 2022, as well as data on income and assets, access to care, usual source of care, preventive care, COVID-19, beneficiary knowledge and information needs, drug coverage, and chronic pain information collected through the Winter and Summer 2023 rounds.⁸ 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



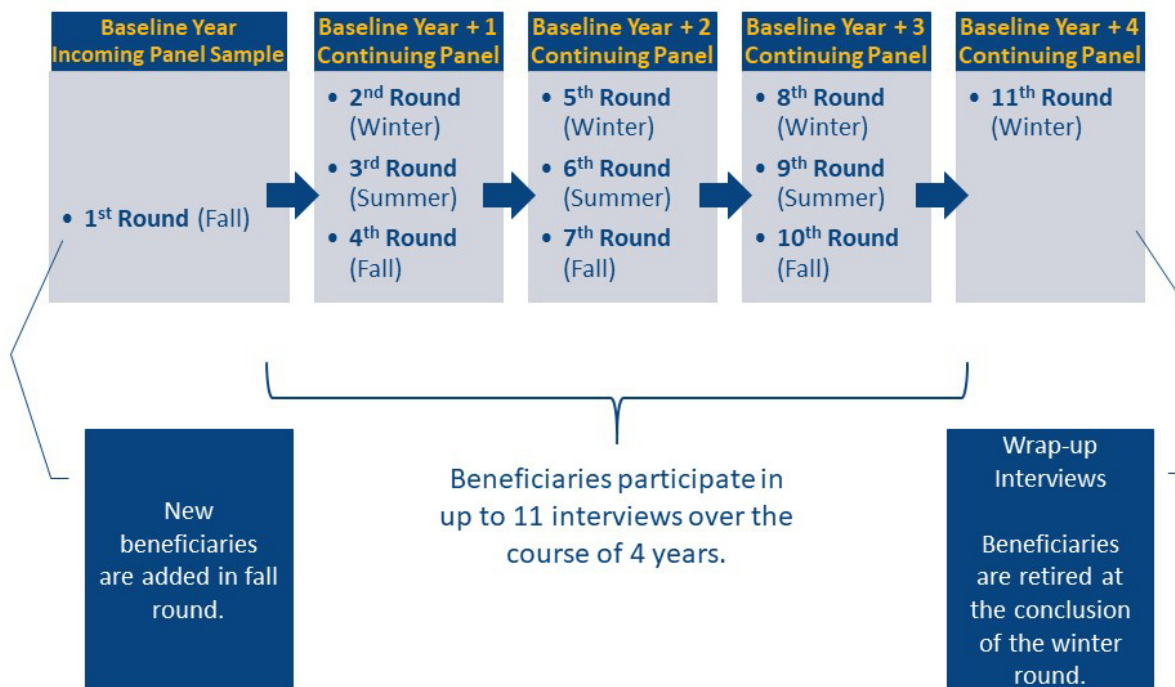
Initial interviews of newly selected beneficiaries take place in the fall round. Since 2016, the fall round begins early in late July or early August to allow more time to conduct outreach and collect information from the new survey respondents who are selected to participate in the MCBS. That is, the early start of the fall round overlaps with the final weeks of data collection for the summer round. These small overlap periods as one round ends and another begins are acceptable design features of the survey.

Subsequent rounds, which occur every four months, involve re-interviewing of the same beneficiary (or appropriate proxy respondents or Facility staff) until they have completed four years of participation (up to 11 interviews in total). Interviews are conducted regardless of whether the beneficiary resides at home or in a long-term care facility, using a questionnaire version appropriate to the setting. Exhibit 4.1.2 depicts the timeline of participation for beneficiaries selected to be in the MCBS sample and Appendix A provides a list of all rounds by data collection year. In Winter 2022 and Summer 2022, physical measures were incorporated

⁸ Due to the nature of some survey items, LDS 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. Please refer to Exhibits 5.2.3 and 5.2.8 for more information.

into the MCBS via a new questionnaire section, but these pilot data were not released. The Summer 2023 pilot data that correspond to the 2022 data year were added to the 2022 Cost Supplement File LDS; please see the *2022 Cost Supplement File Data User's Guide* for additional information.

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.⁹ 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.¹⁰ 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). The size of the new panel is designed to provide a stable number of beneficiaries across all panels participating in the survey annually. Please see Section 6: Sampling for more information on the sample selection.

⁹ 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.

¹⁰ 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: 2018-2022 MCBS Rotating Panel Design

Data Collection Schedule			Panel				
Calendar Year	Season	Round#	2018	2019	2020	2021	2022
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					
2022	Winter	92					
	Summer	93					
	Fall	94					

4.3 Case Types

MCBS respondents are classified by their phase of participation (i.e., Incoming or Continuing) and interview participation (i.e., Community or Facility), which is determined by residence status. These case types are described below.

4.3.1 Incoming and Continuing Cases

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

4.3.2 Community Interviews and Facility Interviews

Approximately 93 percent of the interviews are held with 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 for beneficiaries living in a facility. Over the course of a four-year period, it is not uncommon for beneficiaries to enter long-term care facilities (e.g., nursing homes) or to go back and forth between the community and a facility setting (these cases are called Crossovers). In order to obtain an accurate representation of the experiences of all Medicare beneficiaries, the MCBS includes beneficiaries wherever they reside, even if they reside in and/or enter a facility for the duration of their four years with the study. The MCBS does not conduct Facility interviews with the beneficiary directly; instead, specially trained Facility interviewers administer the survey to Facility administrative staff.

For more information about MCBS data collection procedures and interviewing, see the *2022 MCBS Methodology Report*.

4.4 Completed Interviews

Exhibit 4.4 lists the number of completed interviews for the Fall 2022 Continuing (2019, 2020, and 2021) and Incoming (2022) Panels by age strata. Under the rotating panel design, the beneficiaries selected in Fall 2018 exited the study at the conclusion of the Winter 2022 round.

Exhibit 4.4: 2022 MCBS Fall Round Completed Interviews: Continuing and Incoming Panels

Age Category as of 12/31/2022	2019 Panel	2020 Panel	2021 Panel	2022 Panel	Total
Under 45 years	163	125	191	446	925
45-64 years	147	248	298	599	1,292
65-69 years	201	321	505	1,107	2,134
70-74 years	397	433	444	948	2,222
75-79 years	322	378	492	1,027	2,219
80-84 years	295	389	510	1,048	2,242
85+ years	374	442	501	1,084	2,401
Total	1,899	2,336	2,941	6,259	13,435

SOURCE: 2022 MCBS Internal Sample Control File

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). The beneficiary's residence status determines which questionnaire component is used and how it is administered. See Exhibit 5.1 for a depiction of the MCBS Questionnaire structure.

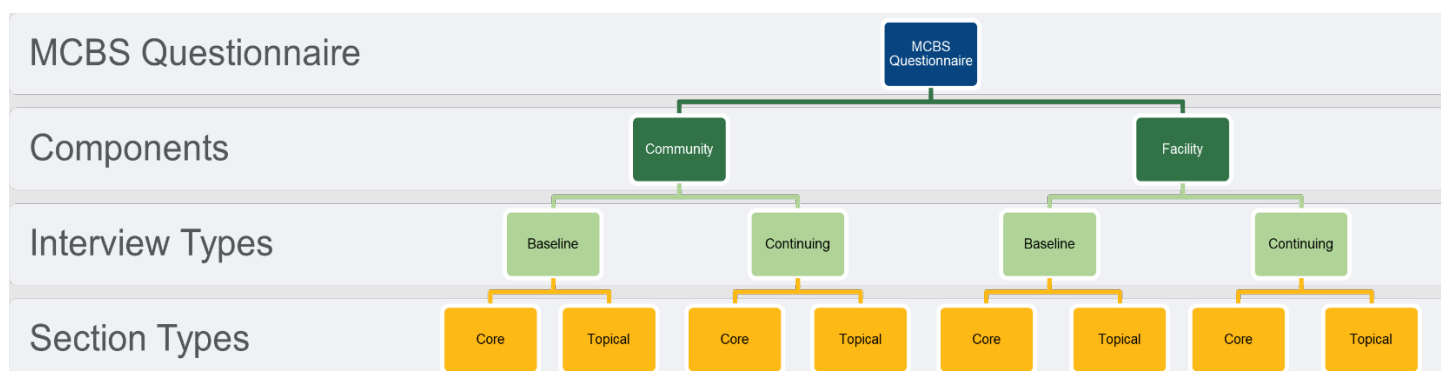
- **Community Component:** Survey administered to beneficiaries living in the community (i.e., not in a long-term care facility such as a nursing home) during the reference period covered by the MCBS interview. An interview may be conducted with the beneficiary or a proxy.
- **Facility Component:** Survey 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. See Sections 5.2 and 5.3 for tables of the 2022 Core and Topical sections.

Exhibit 5.1: MCBS Questionnaire Overview



5.1.1 Items from Validated Scales

The MCBS questionnaire contains content from a variety of sources that are adapted for inclusion in the MCBS. Some questionnaire items on the MCBS come from validated scales that were developed by external researchers and tested for reliability and validity. Two examples of such scales are the Generalized Anxiety Disorder Scale (GAD-2), which is a screening tool for generalized anxiety disorder (see MCBS Community items

HFGAD1 and HFGAD2) and the Patient Health Questionnaire (PHQ-9), which is a screening tool for depression (see MCBS Community items HFPHQ1 through HFPHQ8 and PHQ9QS10).

5.2 Community Questionnaire

The content of the MCBS Community Questionnaire consists of Core and Topical sections. Core survey content is grouped into questionnaire sections that collect data central to the policy goals of CMS. These 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 in each of these categories may be asked each round or seasonally (fall, winter, summer). Data from these questionnaire sections are found in the Survey File and Cost Supplement File data releases. In addition to the Core content, there are several Topical questionnaire sections that capture data on a variety of key topics that relate to the beneficiary's housing characteristics, health behaviors, knowledge about Medicare, and health-related decision making. All data from the Topical sections are included in the Survey File data release.

Different combinations of Core and Topical sections are used depending on a number of criteria, including interview type (Baseline vs. Continuing); the season of data collection (fall, winter, summer); whether the beneficiary is alive, deceased, or in a facility; and whether the interview is being completed with the beneficiary or a proxy.

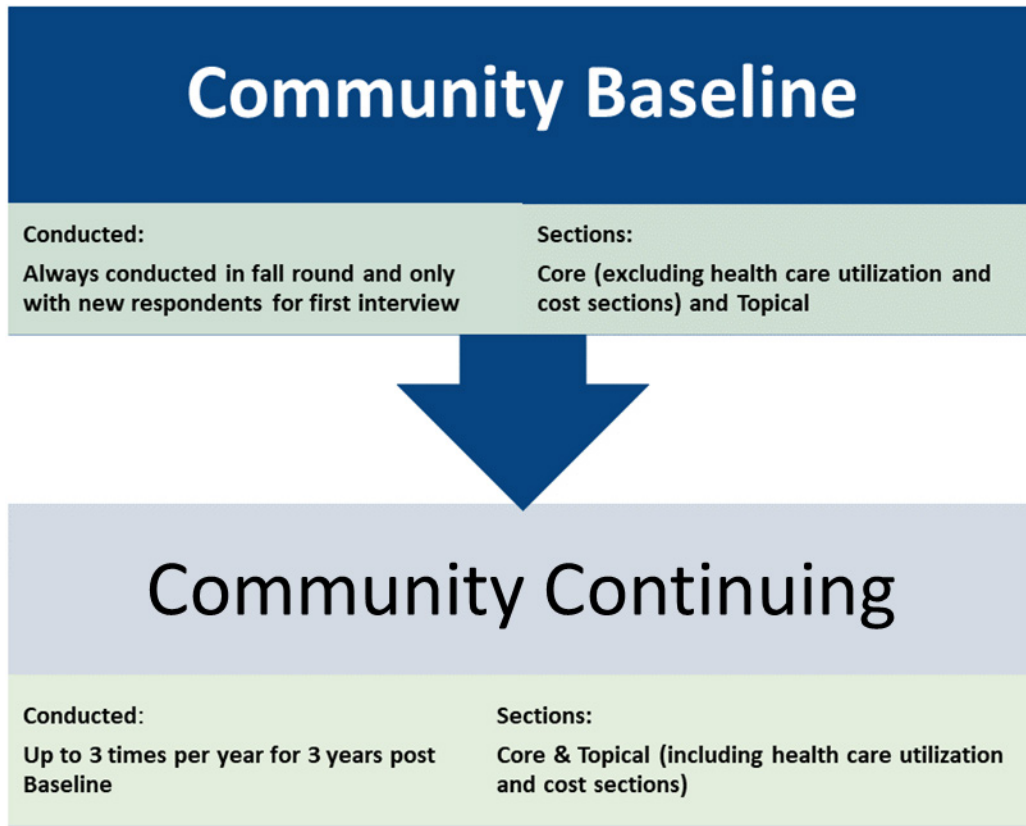
The first Community interview conducted with Incoming Panel respondents is referred to as the Baseline interview. This interview is always conducted in the fall round and consists of a combination of Core and Topical sections. It is important to note that this first interview does not include Core sections that collect health care utilization and cost data. The respondent's 2nd through 11th interviews, also known as the Continuing interviews, consist of Core and Topical sections, including those that collect health care utilization and cost data; these interviews provide three calendar years of reported health care utilization and cost data for each beneficiary.

The Community questionnaire consists of the following components (see Exhibit 5.2):

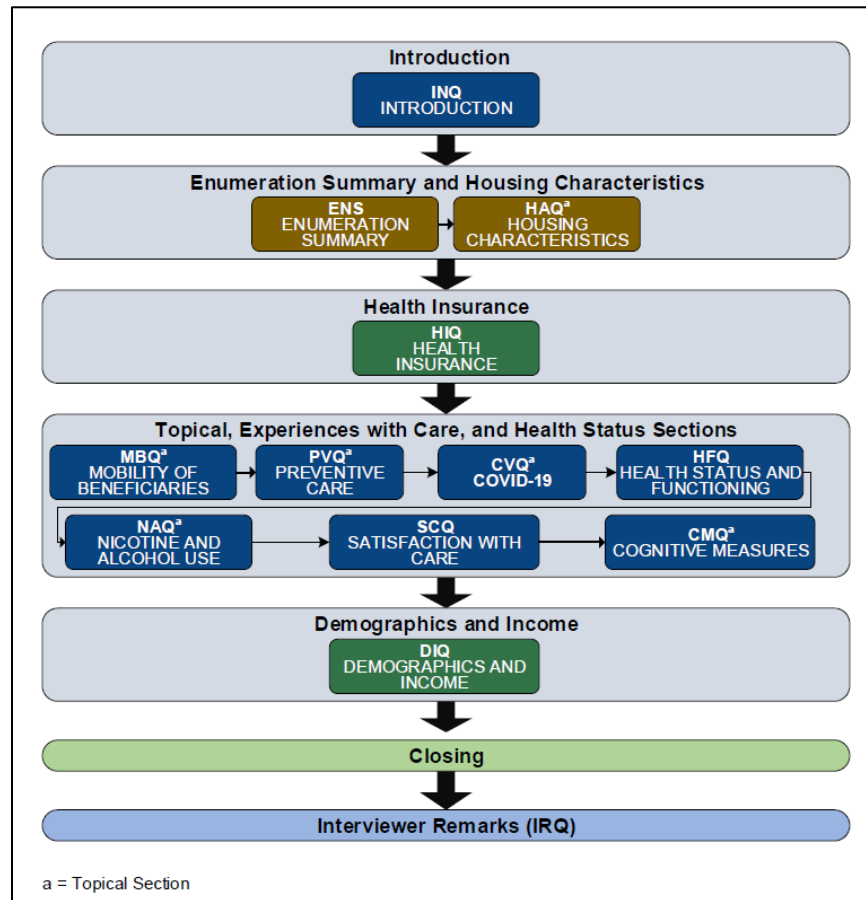
- **Community Baseline questionnaire**
- **Community Continuing questionnaire**

In addition to including data collected in the three rounds (winter, summer, and fall) administered during the calendar year (i.e., January 2022 through December 2022), some data collected in the previous and following calendar years are also included in the 2022 LDS. Specifically, some data collected in 2021 are carried forward to fill in data for 2022 when questionnaire items are administered only once or when data are missing for the data year but valid values exist for the previous year. Some data are also collected in Winter and Summer 2023 and are "pulled back" for inclusion in the 2022 LDS because the section's reference period extends back to 2022; these sections are specified further below.

In this section, data users should note that exhibit titles will indicate either the *data collection year*, which refers to the three rounds (winter, summer, and fall) that occur within the calendar year, or the *data year*, which refers to the data collected over the three years that are included in the LDS.

Exhibit 5.2: Overview of the MCBS Community Questionnaire Components*5.2.1 Baseline Interview*

As the first interview conducted, the Baseline interview provides an opportunity for the field interviewer to develop a strong rapport and connection with the respondent, acquaint the respondent with the intent of the survey, and emphasize the importance of keeping accurate records of medical care and expenses. Whenever possible, field interviewers are assigned to the same beneficiary over the course of their participation in the survey, so establishing a positive relationship is critical during the Baseline interview. Exhibit 5.2.1 depicts the sections and flow of the Community Baseline interview for the 2022 calendar year, which is synonymous with the 2022 data year for Baseline cases.

Exhibit 5.2.1: 2022 Data Collection Year MCBS Community Questionnaire Flow for Baseline Interview

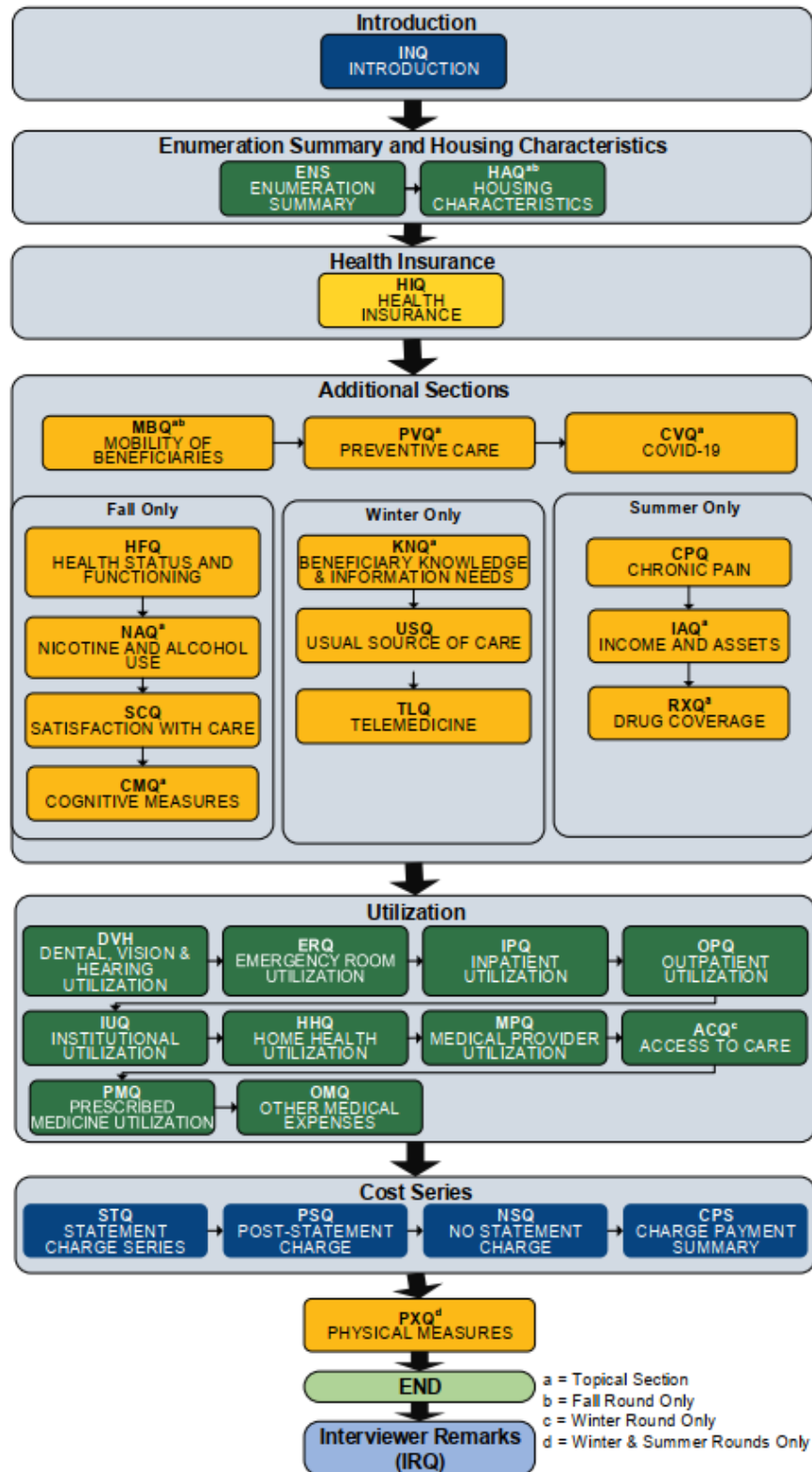
5.2.2 Continuing Interview

The Continuing interview consists of Core sections that focus on the use of medical services and the resulting costs; these sections are asked in essentially the same way each and every time they are administered. The respondent is asked about new health events and to complete any partial information that was collected in the last interview. For example, the respondent may mention a doctor visit during the health care "utilization" part of the interview. In the "cost" section, the field interviewer will ask if there are any receipts or statements from the visit. If the answer is "yes", the field interviewer will record information about the costs from those statements, but if the answer is "no," the question will be stored until the next interview. The Continuing interview also includes sections about health insurance. During each interview, the respondent is asked to verify ongoing health insurance coverage and to report any new health insurance plans.

Continuing interviews also include Topical sections which cover subjects such as mobility or drug coverage. Exhibit 5.2.2 depicts the sections and flow of the Community Continuing interview for the data collection year, rather than the data year, meaning interviews conducted in 2022 used the flow depicted. The data year includes surveys administered in other years which may have slightly different questionnaire flows but are included in the data year LDS given the reference period. This means that some questionnaire sections included in the Survey File will not be reflected in this exhibit if they were first fielded in winter or summer of the following year. To continue to prioritize the collection of seasonal section data and facilitate telephone data collection in response to the COVID-19 pandemic, the questionnaire order was maintained to administer the seasonal sections after Health Insurance (HIQ) and before all utilization and Cost Series sections. This change was made permanent in Summer 2022 due to positive feedback from field interviewers on the revised questionnaire flow.

All sections are considered "Core" sections unless otherwise noted.

Exhibit 5.2.2: 2022 Data Collection Year MCBS Community Questionnaire Flow for Continuing Interview



5.2.3 Core Questionnaire Sections

Each Core section of the Community Questionnaire is described below, organized by topic of information collected. New respondents receiving the Baseline interview do not receive Core sections about health care utilization and costs; these sections are reserved for Continuing respondents. As such, in Fall 2022, only persons in the 2019, 2020, and 2021 Panels received the Core sections about health care utilization and health care costs. All panels received the health insurance section. Exhibit 5.2.3 displays the Core Community questionnaire sections that are included in the Survey File and the Cost Supplement File.

Exhibit 5.2.3: 2022 Data Year MCBS Community Core Sections by Data File and Data Collection Schedule*

Section Group	Abbr.	Section Name	LDS [§]	Data Collection Schedule
Socio-Demographics	IAQ	Income and Assets	SF	Summer 2023**
	DIQ	Demographics/Income	SF	Fall 2022, Baseline Interview
Health Insurance	HIQ	Health Insurance	SF	All Seasons
Utilization	DVH	Dental, Vision and Hearing Care Utilization	CS	All Seasons
	ERQ	Emergency Room Utilization	CS	All Seasons
	IPQ	Inpatient Hospital Utilization	CS	All Seasons
	OPQ	Outpatient Hospital Utilization	CS	All Seasons
	IUQ	Institutional Utilization	CS	All Seasons
	HHQ	Home Health Utilization	CS	All Seasons
	MPQ	Medical Provider Utilization	CS	All Seasons
	OMQ	Other Medical Expenses Utilization	CS	All Seasons
	PMQ	Prescribed Medicine Utilization	CS	All Seasons
Cost	STQ	Statement Cost Series	CS	All Seasons
	PSQ	Post-Statement Charge	CS	All Seasons
	NSQ	No Statement Charge	CS	All Seasons
	CPS	Charge Payment Summary [‡]	CS	All Seasons
Experiences with Care	ACQ	Access to Care	SF	Winter 2023**
	SCQ	Satisfaction with Care	SF	Fall 2022
	TLQ	Telemedicine	SF	Winter 2023**
	USQ	Usual Source of Care	SF	Winter 2023**
Health Status	HFQ	Health Status and Functioning	SF	Fall 2022
	CMQ	Cognitive Measures	SF	Fall 2022
	PXQ	Physical Measures	SF	Summer 2023

SOURCE: MCBS Community Questionnaire

*Certain procedural or operational management sections are collected specifically to manage the data collection process. These sections are not directly included in the LDS files (e.g., Introduction (INQ), Enumeration (ENS), and Interview Remarks (IRQ)).

**These sections are administered in the summer or winter rounds following the current data year given that the reference period is the prior year and data are included in the prior year data files. For guidance on analyzing data from these sections, see Section 9.4.2.

‡Summary sections: Updates and corrections are collected through the summary sections. The respondent is asked to verify summary information gathered in previous interviews. Changes are recorded if the respondent reports information that differs from what was previously recorded.

§Limited Data Set (LDS) indicates the file where the questionnaire data appears (i.e., SF = Survey File, CS = Cost Supplement File).

Socio-Demographics

Two sections in the Community Questionnaire capture key socio-demographic characteristics of the beneficiary.

The **Demographics and Income (DIQ)** section includes traditional demographic items such as Hispanic origin, race, English proficiency, education, and total household income. This section is administered during the Baseline interview.

Income and Assets (IAQ) is a summer round section that collects detailed information about income and assets of the beneficiary and spouse or partner (if applicable). IAQ covers beneficiary (and spouse/partner) income from employment, Social Security, Veteran's Administration, and pensions. The respondent is also asked to indicate the value of the beneficiary's (and spouse's/partner's) assets including retirement accounts, stocks, bonds, mutual funds, savings accounts, businesses, land or rental properties, and automobiles. Also included in this section are items about homeownership or rental status and food security. IAQ collects information about the previous calendar year; thus, income and assets information collected in Summer 2023 (for the 2022 calendar year) are included in the 2022 LDS.

Health Insurance

The Community Questionnaire captures health insurance information each round.

Health Insurance (HIQ) records all health insurance plans that the beneficiary has had since the beginning of the reference period. The survey prompts for coverage under each of the following types of plans: Medicare Advantage, Medicaid, TRICARE, non-Medicaid public plans, Medicare Prescription Drug Plans, and private (e.g., Medigap or supplemental) insurance plans. Detailed questions about coverage, costs, and payment are included for Medicare Advantage, Medicare Prescription Drug, and private insurance plans.

Utilization

The utilization sections of the questionnaire capture health care use by category. Generally, four types of health care utilization are recorded: provider service visits, home health care, other medical expenses, and prescribed medicines. Provider service visits include visits to dental, hearing, and vision care providers; emergency rooms; inpatient and outpatient hospital departments; institutional stays; and medical providers. In these sections, visits are reported as unique events by date, although in cases where there are more than five visits to a single provider during the reference period, the events are entered by month with the number of visits specified. A slightly different reporting structure is used for home health care, other medical expenses, and prescribed medicines.

All utilization sections are administered in all Community Continuing interviews; these sections are not part of the Incoming Panel's Baseline interview. Additional detail is provided on each of the four types of health care utilization collected by the Community Questionnaire below.

Provider Service Visits

The utilization sections collecting provider service dates are as follows.

Dental, Vision, and Hearing Care Utilization (DVH) collects information about dental, vision, and hearing care visits during the reference period as well as other medical expenses such purchases or repairs of glasses and hearing devices. DVH collects the name and type of dental, vision, and hearing care providers; dates of visits; services performed; medicines prescribed during the visits; and any purchases or repairs of glasses and hearing devices. This section replaced the Dental Utilization Questionnaire (DUQ) section from 2018 and earlier.

Emergency Room Utilization (ERQ) records visits to hospital emergency rooms during the reference period. ERQ collects the names of the hospitals, dates of visits, whether the visit was associated with a particular condition, and medicines prescribed during the visits. If a reported emergency department visit resulted in hospital admission, an inpatient visit event is created, with follow up questions asked in the Inpatient Utilization section.

Inpatient Utilization (IPQ) collects information about inpatient stays during the reference period. IPQ collects the names of the hospitals, beginning and end dates of the stays, whether surgery was performed, whether the visit was associated with a particular condition, and medicines prescribed to be filled upon discharge from the hospital (medicines administered during the stay are not listed separately). Inpatient stays resulting from emergency room admissions are also covered.

Outpatient Hospital Utilization (OPQ) prompts for visits that the beneficiary may have made to hospital outpatient departments or clinics during the reference period. OPQ collects the name of the outpatient facility, dates of visits, whether surgery was performed, whether the visit was associated with a particular condition, and medicines prescribed during the visits.

Institutional Utilization (IUQ) collects information about stays in nursing homes or any similar facility during the reference period. IUQ collects the name of the institution(s) and the dates the beneficiary was admitted and discharged from the institution(s).

Medical Provider Utilization (MPQ) collects information about medical provider visits during the reference period. In addition to physicians and primary care providers, this includes visits with health practitioners that are not medical doctors (acupuncturists, chiropractors, podiatrists, homeopaths, naturopaths), mental health professionals, therapists (including speech, respiratory, occupational, and physical therapists), and other medical persons (nurses, nurse practitioners, paramedics, and physician's assistants). MPQ collects names and types of providers, dates, whether the visit is associated with a particular condition, and medicines prescribed during the visit.

Home Health Care Visits

Home Health Utilization (HHQ) collects information about home health provider visits from both professional and non-professional providers, during the reference period. HHQ collects names and types of home health providers, dates of visits, and services performed during visits.

Prescribed Medicines

The **Prescribed Medicine Utilization (PMQ)** section collects details about prescribed medicines obtained during the reference period. For medicines recorded in the provider service visit sections (in the context of those visits), PMQ collects the medicine strength, form, quantity, and number of purchases. Medicines that were not previously reported during the course of the provider service visit utilization sections, including those that are refilled or called in by phone, are also collected in this section. Unlike for provider service visits, event dates are not collected for prescribed medicines. Instead, the interviewer records the number of purchases or refills. Information is not collected about non-prescription medicines and prescriptions that are not filled.

Other Medical Expenses

The Community Questionnaire also records other medical expenses. These expenses are reported using a slightly different reporting structure within the questionnaire. The reporting structure used to capture other medical expenses within the questionnaire differs slightly than that used for capturing provider services events. For example, as opposed to capturing details about a visit to a provider (e.g., provider name, date of visit, etc.), the questionnaire records the date(s) the beneficiary rented, purchased, or repaired each type of medical equipment.

Other Medical Expenses Utilization (OMQ) collects information about medical equipment and other items (excluding prescriptions) that the beneficiary purchased, rented, or repaired during the reference period. Other medical expenses include orthopedic items (wheelchairs, canes, etc.), diabetic equipment and supplies, dialysis equipment, prosthetics, oxygen-related equipment and supplies, ambulance services, other medical equipment (beds, chairs, disposable items, etc.) and alterations to the home or car. For each item, the date(s) of rental, purchase, or repair are recorded. For disposable medical items (e.g., bandages), the number of purchases is collected, rather than a date.

Data collected in the utilization sections are released with the Cost Supplement File LDS. See the *Cost Supplement File: Data User's Guide* for more information.

Cost Series

Once all utilization sections are completed, the questionnaire flows to the cost series, wherein the costs of all reported visits and purchases are recorded, along with the amount paid by various sources. Importantly, additional visits and purchases not reported in the utilization sections of the questionnaire could be recorded within the cost series, and all corresponding data for those events are collected within the cost series.

The cost series consists of four sections: Statement, Post-Statement, No Statement, and Charge Payment Summary. Each is summarized in Exhibit 5.2.4 and described below.

Exhibit 5.2.4: Cost Series Section Overview

Statement Series (STQ)	Post-Statement Series (PSQ)
Collect cost information from: <ul style="list-style-type: none"> • Medicare • Insurance • TRICARE • Drug plan statements 	Collect costs for "rent-to-buy" items <ul style="list-style-type: none"> • Only administered to a small percentage of respondents
No Statement Series (NSQ)	Charge Payment Summary (CPS)
Collect information from: <ul style="list-style-type: none"> • Bills • Receipts • Invoices 	Collect information on outstanding charges from: <ul style="list-style-type: none"> • Statement paperwork • Non-statement paperwork

The **Statement Cost Series section (STQ)** collects medical cost information directly from Medicare Summary Notices (MSNs), insurance explanations of benefits (EOB), Prescription Drug Plan statements, and TRICARE or other insurance statements. In cases where the beneficiary had more than one payer (e.g., Medicare and private insurance), interviewers organize statements into charge bundles, which are driven by the claim total on an MSN or EOB and may include one or more utilization events (visits, medicines, or purchases). Each charge bundle is entered separately, and all previously reported events associated with the charge bundle are linked to the cost record. Payment details are entered from the statements and any remaining amount not accounted for is confirmed with the respondent. This process is repeated for all

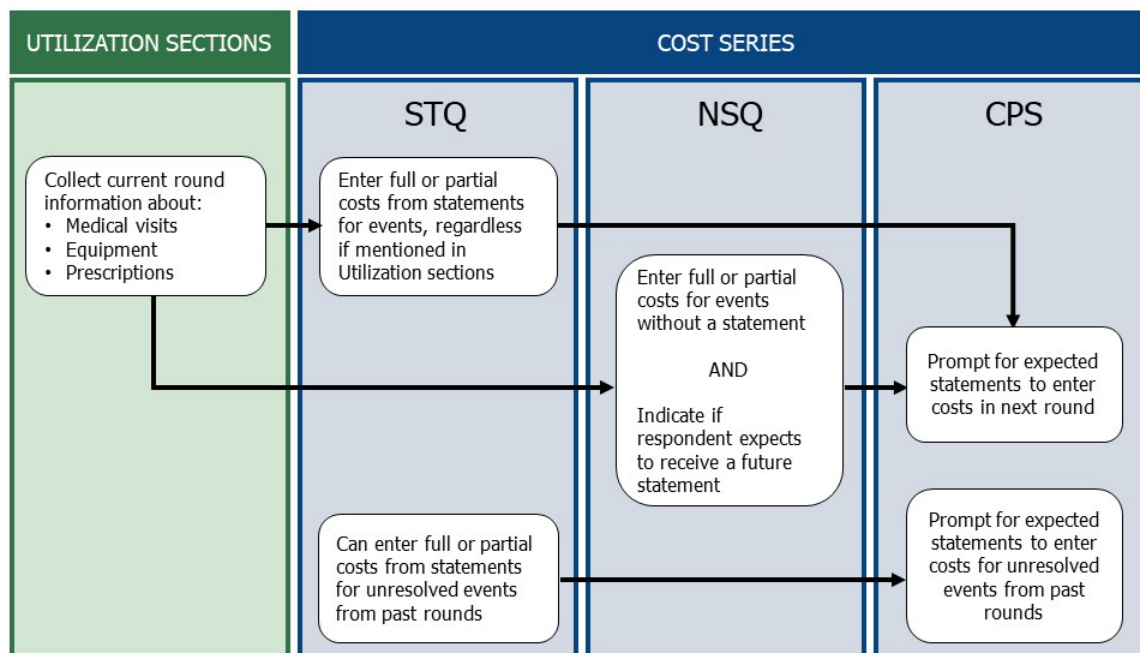
available, not previously recorded insurance statements containing events that occurred within the survey reference period (roughly the past year).

The **Post-Statement Charge section (PSQ)** facilitates cost data collection for rental items that span multiple rounds of interviews (such as a long-term wheelchair rental) and for which cost data has not yet been reported.

The **No Statement Charge section (NSQ)** prompts for cost data for all events that do not have a Medicare, insurance, or TRICARE statement reported in the current round. This section attempts to capture cost data even in absence of insurance statements. The respondent may refer to non-statement paperwork such as bills or receipts to help collect accurate cost information. NSQ loops through a series of cost verification items for each event or purchase reported during the current round utilization but not already linked to a cost record via the Statement section. If respondents indicate a statement for the event is expected, then the NSQ items are bypassed.

The final cost series section, the **Charge Payment Summary (CPS)**, reviews outstanding cost information reported within the last two rounds. For example, if the respondent reported in the previous interview that he/she expected to receive an insurance statement for a particular event, then this event is carried forward to CPS in the next round. Any charge bundle for which costs are not fully resolved is asked about in the next round's CPS section. There are a variety of reasons a cost record might qualify to be asked about in CPS (referred to as "CPS Reasons"). For example, a respondent may have been expecting to receive a statement related to the event or may have reported payments that account for only part of the total charge. The amount of information collected in CPS and the path through the section is determined by the CPS reason for the cost record. One case can have multiple cost records flagged for CPS with a variety of CPS reasons. The questionnaire loops through each eligible cost record in an attempt to collect further cost data.






The flow of sections and questions within the cost series varies depending on data collected in the current round (e.g., whether the beneficiary had a health insurance statement for a visit reported in the current round) and data collected in prior rounds (i.e., whether there was outstanding cost information reported from a prior round). Exhibit 5.2.5 illustrates how paths through these sections may vary depending on health care utilization and cost information collected in the current and previous rounds.

Exhibit 5.2.5: Utilization and Cost Section Flow

Costs are considered unresolved when full cost information is not collected due to events being reported 1) without any cost or payment information, 2) with an indication that a statement is expected, so follow-up questions about costs and payments are deferred until the next interview, or 3) with partial information about costs or payments, but there is a remaining dollar amount with pending payment information.

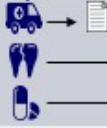





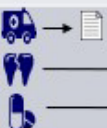


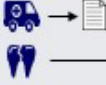

The current MCBS protocol allows for cost resolution attempts up to two rounds later than the events were reported. Exhibit 5.2.6 displays sample paths to resolving cost information. The first row displays a hospital event reported with costs and statement. This cost is resolved within the round. The second row displays a resolved dental event reported in the summer round with the statement provided in the fall round. The third row displays a prescription medicine event reported with a statement in the summer and resolved in the winter after the statement was provided. The final row displays an unresolved event that was reported in the summer round but did not receive cost or statement verification.

Exhibit 5.2.6: Example Paths Toward Cost Resolution

Scenario	Summer Round	Fall Round	Winter Round	Cost Status
Event reported with costs, statement available				Resolved
Event reported without costs, awaiting statement				Resolved
Event reported with receipt, awaiting statement				Resolved
Event reported without costs, statement not received				Unresolved

The 2022 data year includes 2022 events collected from Winter 2022 through Summer 2023 (see Exhibit 5.2.7). The unresolved costs are indicated with a red circle-backslash symbol and are unresolved given that the statement was not received.

Exhibit 5.2.7: Events Collected in the 2022 Data Year

Winter 2022 (R92)	Summer 2022 (R93)	Fall 2022 (R94)	Winter 2023 (R95)	Summer 2023 (R96)
				
				
				
				

Data collected in the cost series are released with the Cost Supplement File LDS. See the *Cost Supplement File: Data User's Guide* for more information.

Experiences with Care

Four sections cover the beneficiary's experience with care in various medical settings.

Access to Care (ACQ) is administered in the winter round for Continuing respondents and focuses on the beneficiary's experience with particular types of medical encounters (hospital emergency room, hospital clinic or outpatient department, long-term care facility, or medical doctor visits) during the reference period. If the beneficiary had one or more of a particular type of medical encounter, additional items collect information about services received and waiting times associated with the most recent encounter. ACQ collects information about the previous calendar year; thus, ACQ data collected in Winter 2023 (for the 2022 calendar year) are included in the 2022 LDS.

Satisfaction with Care (SCQ) is part of the fall round interview for Incoming Panel and Continuing respondents and collects the respondent's opinions about the health care that the beneficiary had received. The questions refer to medical care received from all medical providers, including both doctors and hospitals.

The **Telemedicine (TLQ)** section is administered in the winter round interview for Continuing respondents who report a usual source of care in the USQ section. TLQ asks questions on the availability and utilization of telemedicine services. TLQ collects information about the previous calendar year; thus, TLQ data collected in Winter 2023 (for the 2022 calendar year) are included in the 2022 LDS.

The **Usual Source of Care (USQ)** section is administered in the winter round for Continuing respondents and collects specific information about the usual source of health care for the beneficiary as well as any

specialists seen during the reference period. USQ collects information about the previous calendar year; thus, USQ data collected in Winter 2023 (for the 2022 calendar year) are included in the 2022 LDS.

Health Status

Health Status and Functioning (HFQ) collects information on the beneficiary's general health status and needs. This includes specific health areas such as disabilities, vision, hearing, and preventive health measures. HFQ includes measures of the beneficiary's ability to perform physical activities, moderate and vigorous exercise, health care maintenance and needs, and standard measures of Instrumental Activities of Daily Living (using the telephone, preparing meals, etc.), and Activities of Daily Living (bathing, walking, etc.). In addition, HFQ asks about medical diagnoses for common conditions (cancer, arthritis, hypertension, etc.). Finally, the section covers mental health conditions, falls, urinary incontinence, and a more extensive series of questions for beneficiaries with high blood pressure and diabetes.

Cognitive Measures (CMQ) contains four well-established cognitive measures to assess cognitive functioning among beneficiaries, including backwards counting, date naming, object naming, and president/vice president naming.

The **Physical Measures Questionnaire (PXQ)** section collects six measures: gait speed, chair stand, balance test, measured height, measured weight, and measured grip strength. The PXQ section is administered in the summer round to Continuing respondents. The section appears at the end of the interview and only during interviews conducted in person with the beneficiary.

Operational and Procedural

These sections help guide the interviewer through the interview, providing scripts for introducing and ending the interview. They also facilitate collection of information about household members to augment sample information. Data collected in these sections are not included in the Survey or Cost Supplement data files.

Introduction (INQ) introduces the survey and records whether the interview was completed by the beneficiary or a proxy. For interviews completed by a proxy, the introduction collects the proxy's name and relationship to the beneficiary and determines if the proxy is a member of the beneficiary's household. The introduction is part of every Community Questionnaire.

The **Closing (END)** section is administered to close the interview for all respondents. During the exit interview, this section contains additional scripts to thank the respondent for participation over the four years of the MCBS.

Enumeration (ENS) collects household information and a roster of persons living in the household. For each household member added to the roster, his/her relationship to the beneficiary, sex, date of birth, age, and employment status are collected. ENS is administered in all rounds except the final exit interview.

The **Interviewer Remarks Questionnaire (IRQ)** captures additional metadata about the interview, as recorded by the interviewer. This includes the length of the interview, assistance the respondent may have received, perceived reliability of the information provided during the interview, and comments the interviewer had about the interviewing situation. IRQ is administered after every interview, but it is generally completed after leaving the respondent's home, as none of the questions are directed to the respondent.

5.2.4 Topical Questionnaire Sections

Each Topical section is described below, organized by type of information collected. Exhibit 5.2.8 lists the Topical sections and data collection season. Note that information collected via Topical questionnaire sections is included in the Survey File only and is not included in the Cost Supplement File. In addition, some Topical questionnaire section data are collected through the summer following the current data year (i.e., IAQ, KNQ, PVQ, CPQ, CVQ, and RXQ). Annually, special non-response adjustment weights are included within certain Survey File segments for use in analysis when data are not collected within the same calendar year. Most but not all Topical questionnaire sections have corresponding Topical segments with these special non-response adjustment weights (see Exhibits 9.4.1 and 9.4.2 for more information).

Exhibit 5.2.8: 2022 Data Year MCBS Community Topical Sections by Data File and Data Collection Schedule

Section Group	Abbr.	Section Name	LDS*	Data Collection Schedule
Housing Characteristics	HAQ	Housing Characteristics	SF	Fall 2022
Social Determinants of Health or Health Behaviors	CPQ	Chronic Pain	SF	Summer 2023
	MBQ	Mobility of Beneficiaries	SF	Fall 2022
	NAQ	Nicotine and Alcohol Use	SF	Fall 2022
	PVQ	Preventive Care	SF	Fall 2022, Winter 2023, and Summer 2023 [±]
	IAQ	Food Insecurity items	SF	Summer 2023 ^{**±}
COVID-19	CVQ	COVID-19	SF	Fall 2022, Winter 2023, and Summer 2023
Knowledge and Decision Making	KNQ	Beneficiary Knowledge and Information Needs	SF	Winter 2023 [±]
	RXQ	Drug Coverage	SF	Summer 2023 [±]

SOURCE: MCBS Community Questionnaire

*LDS indicates the file where the questionnaire data appears (i.e., SF = Survey File, CS = Cost Supplement File).

**The Food Insecurity items are included within the Income and Assets Questionnaire (IAQ).

[±]Section is administered in a round following the current data year. The reference period for this section is the prior year and data are included in the prior year data files. For guidance on analyzing data from these sections, see Section 9.4.2.

Housing Characteristics

Housing Characteristics (HAQ) collects information on the beneficiary's housing situation. This includes the type of dwelling, facilities available in the household (e.g., kitchen and bathrooms), accessibility, and modifications to the home (e.g., ramps, railings, and bathroom modifications). This section also records if the beneficiary lives in an independent or assisted living community (distinct from a nursing or long-term care facility) where services like meals, transportation, and laundry may be provided. HAQ is administered in the fall for all beneficiaries in the Community component.

Social Determinants of Health or Health Behaviors

Some questionnaire sections record additional information about health behaviors, specifically mobility, preventive care, and nicotine and alcohol use.

Chronic Pain (CPQ) is a summer round section that collects information about beneficiaries' experiences with chronic pain. The CPQ begins with PAINOFTN, which asks whether or not beneficiaries experienced pain

within the last three months. If so, the section asks more detailed questions about the beneficiary's experience with pain and what types of services and activities they have used to manage their pain. The CPQ section is not administered to proxy respondents. Questionnaire items were developed by the National Pain Strategy (NPS) Population Research Working Group for inclusion in federal surveys.

Mobility of Beneficiaries (MBQ) is a fall round section that determines the beneficiary's use of available transportation options, with a focus on reduced mobility and increased reliance on others for transportation.

The **Preventive Care (PVQ)** section collects information about beneficiaries' preventive health behaviors. Questions administered in this section vary by data collection season. In the winter round, the PVQ focuses on the influenza vaccine, while in the summer round, the PVQ asks about the shingles and pneumonia vaccines. In the fall round, the PVQ asks whether the beneficiary has received various types of applicable preventive screenings or tests, such as a mammogram, Pap smear, or digital rectum exam. In the summer and winter rounds, PVQ collects information about the previous calendar year; thus, PVQ data collected in Winter and Summer 2023 (for the 2022 calendar year) are included in the 2022 LDS.

Income and Assets (IAQ) is a summer round section that collects detailed information about income and assets of the beneficiary and spouse or partner (if applicable), however it also includes items about food security. IAQ collects information about the previous calendar year; thus, food insecurity data collected in Summer 2023 (for the 2022 calendar year) are included in the 2022 LDS.

Nicotine and Alcohol Use (NAQ) collects information on beneficiaries' smoking behavior, including past and current use of cigarettes, cigars, "smokeless" tobacco, and e-cigarettes. It also asks about past and current drinking behavior.

COVID-19

The **COVID-19 (CVQ)** section collects vital information on how the Medicare population is impacted by the COVID-19 pandemic. CVQ spans a number of COVID-related topics, including presumptive vaccine uptake and vaccine utilization.

Knowledge and Decision-Making

Respondent's knowledge of Medicare and health-related decision-making is captured in the following Topical sections.

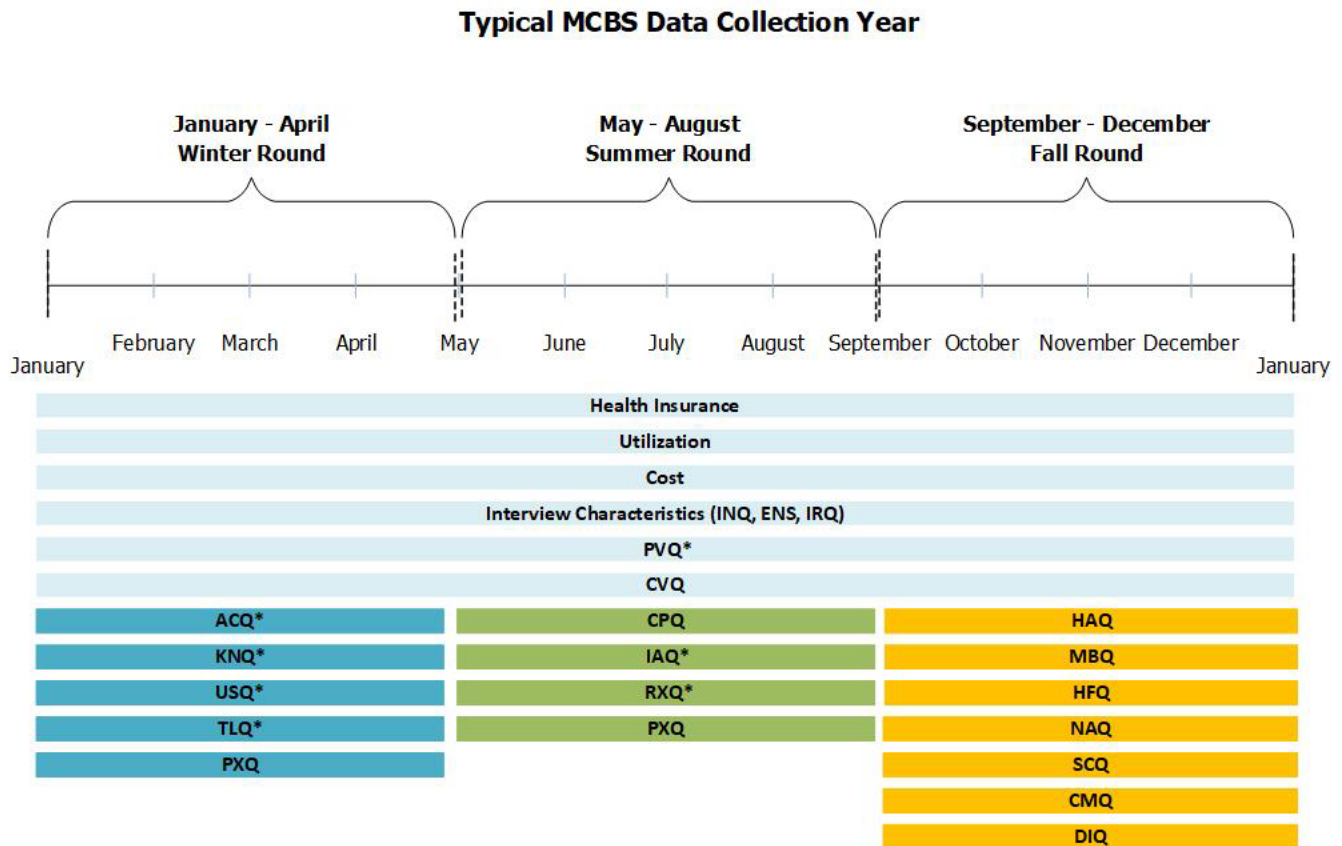
The **Beneficiary Knowledge and Information Needs (KNQ)** section is administered in the winter round. These items measure the respondent's self-reported understanding of Medicare and common sources of information about health care and Medicare. KNQ collects information about the previous calendar year; thus, KNQ data collected in Winter 2023 (for the 2022 calendar year) are included in the 2022 LDS.

The **Drug Coverage (RXQ)** section is a summer round section that focuses on the Medicare Prescription Drug benefit, including respondent knowledge of the benefit, and opinions of the beneficiary's drug coverage, whether through a Medicare Prescription Drug Plan, a Medicare Advantage plan with prescription drug coverage, or a private insurance plan that covers prescription drugs. RXQ collects information about the previous calendar year; thus, RXQ data collected in Winter 2023 (for the 2022 calendar year) are included in the 2022 LDS.

5.2.5 Community Questionnaire Section Rotation within a Data Year

Exhibit 5.2.9 presents the MCBS Questionnaire section rotation schedule for 2022. The 2022 MCBS data releases reflect data collected from January 2022 through December 2022 and also includes data collected in Winter and Summer 2023 rounds from questionnaire sections with a 2022 reference period.

Exhibit 5.2.9: 2022 Data Collection Year MCBS Community Questionnaire Section Rotation



*Fielded in 2023, but given the reference period is 2022, data are included in the 2022 LDS.

5.3 Facility Instrument

In addition to collecting information from respondents living in the community, the MCBS collects information at the institutional level if the beneficiary is living in a facility at the time of the interview. Information is obtained only by interviewing Facility staff; the beneficiary is never interviewed directly.

Similar to the Community Questionnaire, if a beneficiary is living in a facility when first selected to participate in the MCBS, a Facility Baseline interview is administered. For cases in the 2nd through 11th round, a Facility Continuing interview is conducted. While administration of the Facility Instrument sections varies by season and interview type, the Facility Instrument is comprised exclusively of Core sections; each section collects information that is considered of critical importance to the MCBS.

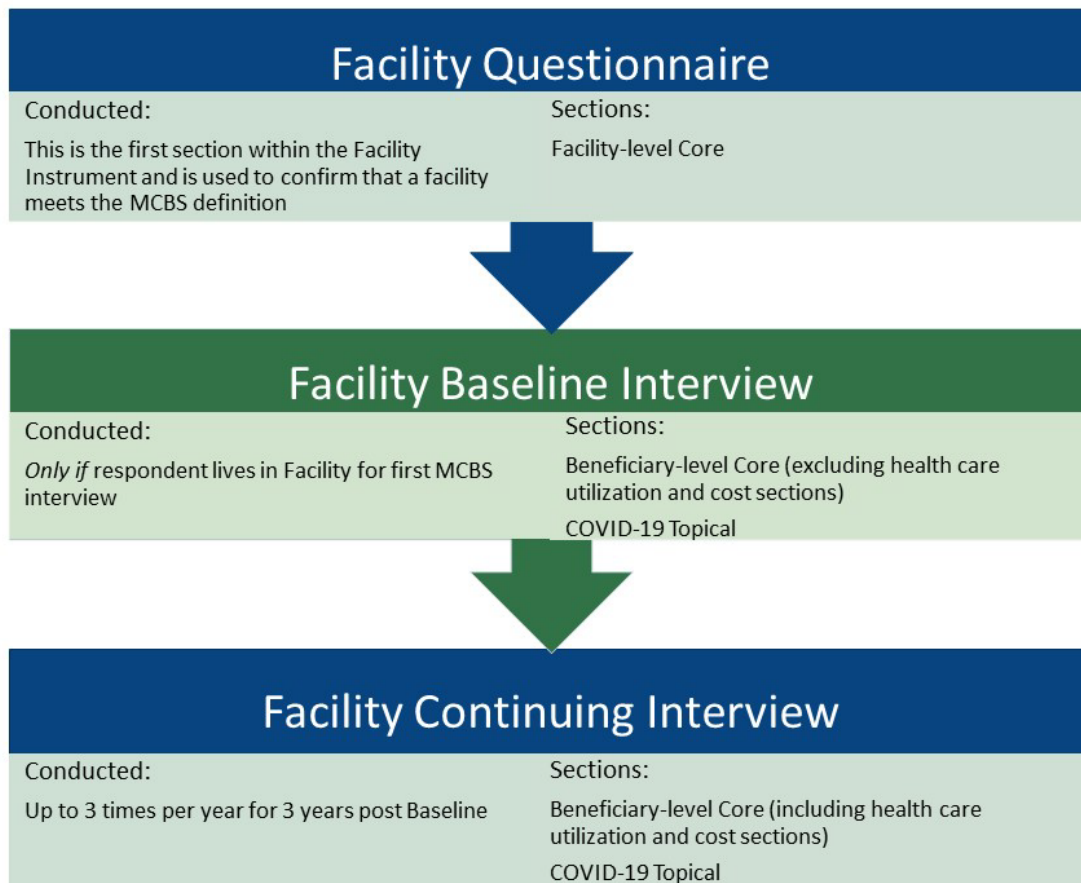
The Facility Instrument consists of the following components (see Exhibit 5.3):

- Facility Questionnaire
- Facility Baseline interview
- Facility Continuing interview

Due to the redesign of the MCBS Facility Instrument in Fall 2019, the instrument flow varies for Medicare and/or Medicaid-certified facilities and facilities not certified by Medicare and/or Medicaid. Facilities that report a CMS Certification Number (CCN) and are therefore certified by Medicare and/or Medicaid receive a shortened MCBS Facility Instrument, as the FQ and HS sections skip variables redundant with Minimum Data Set (MDS) and Certification and Survey Provider Enhanced Reports (CASPER) administrative data. Variables skipped during interview administration are instead populated using MDS and CASPER administrative data sources during data processing. Facilities that do not report a CCN receive the full MCBS Facility Instrument.

If a person living in a facility returns to the community, that person would receive the Community Questionnaire. If the beneficiary spent part of the reference period in the community and part in a facility, then a separate interview is conducted to collect information pertaining to the beneficiary's experiences covering each distinct period of time. In this way, a beneficiary is followed in and out of facilities and a continuous record is maintained regardless of the location of the beneficiary.

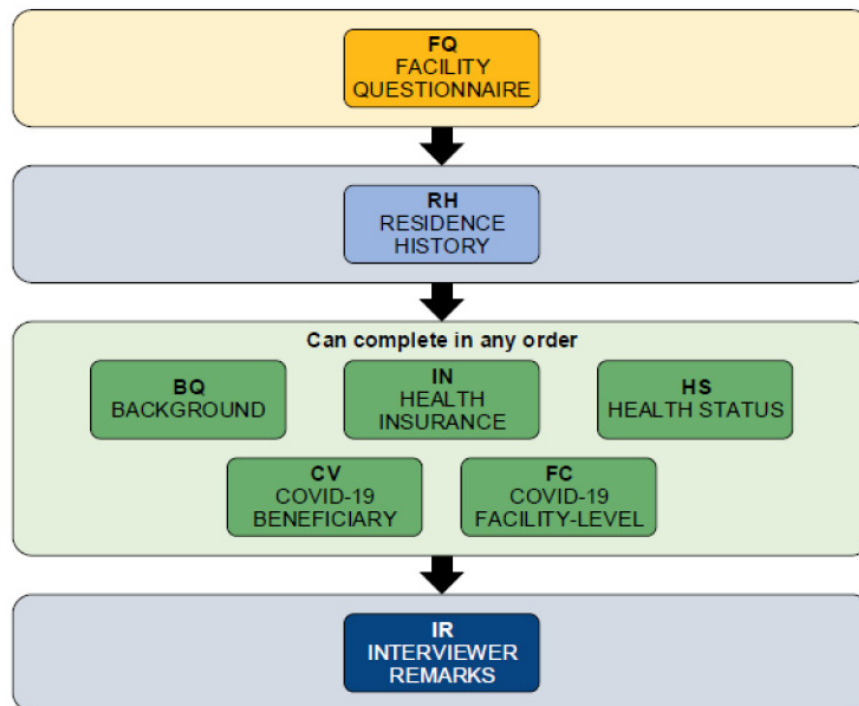
Exhibit 5.3: Overview of the MCBS Facility Instrument



5.3.1 Facility Baseline Interview

The Facility Baseline interview (see Exhibit 5.3.1) serves as a reference interview and gathers information on the facility itself as well as the health status, insurance coverage, residence history, and demographic information for the beneficiary. This flow depicts the sections and flow of the Facility Baseline interview for the 2022 calendar year, which is synonymous with the 2022 data year for Baseline cases.

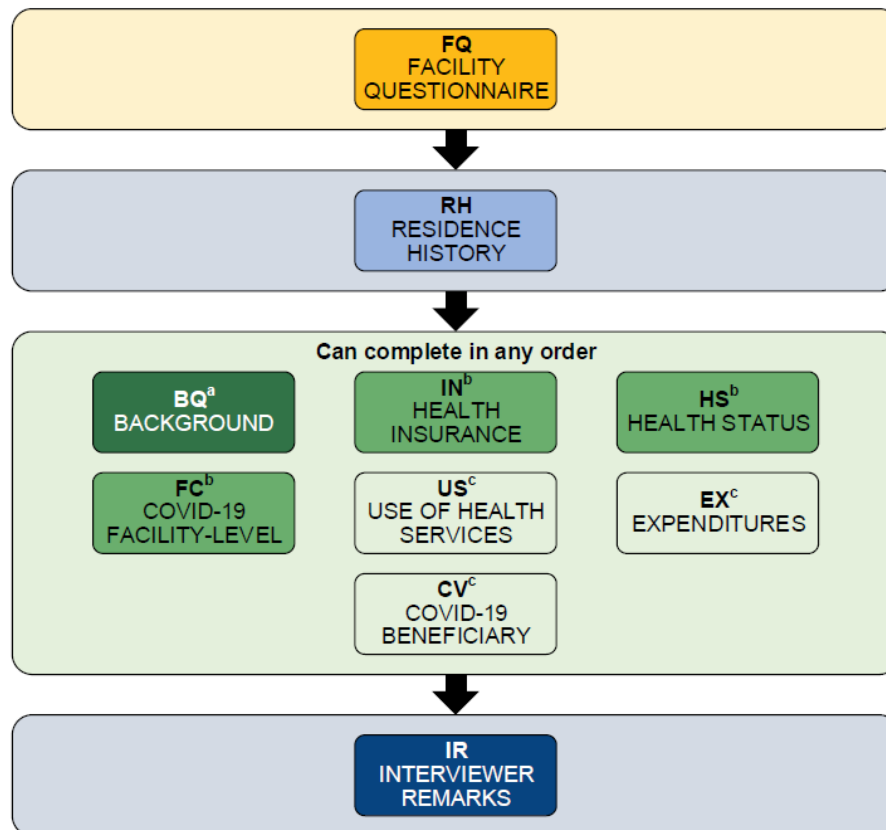
Exhibit 5.3.1: 2022 Data Collection Year MCBS Facility Instrument Flow for Baseline Interview



5.3.2 Facility Continuing Interview

Exhibit 5.3.2 illustrates the flow of the Facility Continuing interview sections. This flow reflects the data collection year, rather than the data year, meaning the interviews conducted in the 2022 calendar year used the flows depicted. Note that beneficiaries who move to a facility from the community (Community to Facility cases), move to a new facility (Facility to Facility cases), or move to the community from the facility (Facility to Community cases) receive a different combination of Facility Continuing sections than beneficiaries who have lived continuously in the same facility.

Exhibit 5.3.2: 2022 Data Collection Year MCBS Facility Instrument Flow for Continuing Interviews



a = Administered only for Community to Facility interviews

b = Administered to all sample types in Fall round. Otherwise, administered only for Community to Facility, Facility to Facility, and for beneficiaries living in a Facility whose last interview was a Community interview and who completed a Facility interview in a prior round.

c = Administered for all Facility interviews

5.3.3 Facility Continuing Core Sections

Each Core section of the Facility Instrument is described below, organized by topic of information collected. The sections depicted in Exhibit 5.3.3 parallel the Core sections for the Community component. These sections of the Facility Continuing interview are administered in the same rotation as the Community Continuing interview (the 2nd through the 11th rounds); however, beneficiaries new to a facility receive additional Core sections.

Similar to the Community Questionnaire, operational management/procedural data are collected through the Interviewer Remarks (IR) section, which is completed by the interviewer and primarily used for case finalization. Exhibit 5.3.3 summarizes each component of the Facility questionnaire by data release.

Exhibit 5.3.3: 2022 Data Year MCBS Facility Core Sections by Data File and Data Collection Schedule*

Section Group	Abbr.	Section Name	LDS [§]	Data Collection Schedule
Facility Characteristics	FQ	Facility Questionnaire	SF	All Seasons
Socio-Demographics	RH	Residence History	SF	All Seasons
	BQ	Background	SF	Fall 2022, Baseline Interview**
Health Insurance	IN	Health Insurance	SF	Fall 2022 [±]
Utilization	US	Use of Health Services	CS	All Seasons
Cost	EX	Expenditures	CS	All Seasons
Health Status	HS	Health Status	SF	Fall 2022 [±]

SOURCE: MCBS Facility Instrument

*Certain procedural or operational management sections are collected specifically to manage the data collection process. These sections are not directly included in the LDS files (e.g., Interviewer Remarks (IR)).

**The BQ section is also administered to Community-to-Facility Crossover cases each season.

[±]The IN and HS sections are also administered to Community-to-Facility and Facility-to-Facility cases each season.

[§]Limited Data Set (LDS) indicates the file where the questionnaire data appears (i.e., SF = Survey File, CS = Cost Supplement File).

Facility Characteristics

The Facility Characteristics Core section contains the **Facility Questionnaire (FQ)** section of the Facility Instrument. The FQ section collects information on the number, classification, and certification status of beds within the facility; sources of payment for Facility residents; and Facility rates. Interviewers typically conduct the FQ with the Facility administrator. Interviewers are not allowed to abstract this section of the interview; it must be conducted with a Facility staff member.

Since the 2019 Facility Instrument redesign and the usage of administrative data for Medicare and/or Medicaid certified facilities, only facilities that do not report a CCN receive the full FQ section.

Socio-Demographics

The Socio-Demographics Core sections capture key characteristics of the interview and the beneficiary. These include residence history and demographics.

The **Residence History (RH)** section collects information about all of the places that the beneficiary stayed during the reference period. Information is collected about where the beneficiary was just before entering the facility and where he/she went if they had been discharged. For each stay, the interviewer collects the name

of the place of residence, the type of place it is, and the start and end date for the period the beneficiary was living there.

The RH section creates a timeline of the beneficiary's whereabouts from the date the beneficiary entered the facility or the date of the last interview, through the date of interview, date of discharge, or date of death. The goal is to obtain a complete picture of the beneficiary's stays during the reference period, including any stays of one night or more in hospitals, other facilities, or any other place.

The **Background Questionnaire (BQ)** collects background information about the beneficiary, such as use of long-term care before admission to the facility, level of education, race, ethnicity, service in the Armed Forces, marital status, spouse's/partner's health status, living children, and income. The BQ is completed only once for each beneficiary during their first interview in the facility.

Health Insurance

The Health Insurance Core section contains the **Health Insurance (IN)** section of the Facility Instrument. The IN section collects information about the beneficiary's type(s) of health insurance coverage. This includes questions about all types of health insurance coverage the beneficiary had in addition to Medicare: private insurance, long-term care insurance, Department of Veterans Affairs eligibility, and TRICARE or CHAMPVA.

Because of differences in interview setting, the content collected in the IN section differs from the content collected in the INQ section of the Community Questionnaire. For example, because the Facility Instrument is administered to Facility staff, as opposed to interviewing the beneficiary directly, the Facility Instrument collects the name of the insurance company for a beneficiary's private insurance plan but does not collect follow-up details about whether the plan was purchased through an employer or some other way.

Utilization

The **Use of Health Care Services (US)** section collects information on the beneficiary's use of health care services while a resident of the facility. This includes visits with a range of providers including medical doctors, dentists, and specialists; visits to the hospital emergency room; and other medical supplies, equipment, and other types of medical services provided to the beneficiary.

The best Facility respondent for this questionnaire section is usually someone directly involved with the beneficiary's care or someone who is familiar with the medical records.

Data collected in US are released with the Cost Supplement File LDS. See the *Cost Supplement File: Data User's Guide* for more information.

Cost

The Facility Cost component consists of the **Expenditures (EX)** section. The EX section collects information about bills for the beneficiary's care at a facility and payments by source for those charges. Data are only collected for the time period when the beneficiary was a resident of the facility at which the interview takes place. The EX section collects information by billing period (e.g., monthly, semi-monthly, quarterly, etc.).

Unlike the Community Questionnaire, which collects cost information for each service, the EX section collects only the fees the facility bills for the beneficiary's care. The EX section collects information on the amount billed for the beneficiary's basic care and for any health related ancillary services. Typically, the EX section is administered to Facility staff located in the billing office.

Data collected in EX are released with the Cost Supplement File LDS. See the *Cost Supplement File: Data User's Guide* for more information.

Health Status

The **Health Status (HS)** section collects information on the beneficiary's general health status, ability to perform various physical activities, general health conditions, IADLs, and ADLs.

Most of the information needed to conduct the HS section may be found in a medical chart. The Federal Government requires that all nursing facilities certified by Medicaid or Medicare conduct comprehensive and standardized assessments of each resident's health status when the resident is admitted to the nursing home and at regular intervals thereafter.¹¹ These assessments are captured by the Long-Term Care MDS,¹² which contains a set of key items measuring a resident's capacity to function independently. Nursing homes use this information to assess each resident's health status, identify problem areas and, where problems exist, formulate care plans to address them.

The HS section is designed to mirror the flow and wording of the MDS items; it contains a subset of the MDS items. In addition, the HS section contains some questions that are not found on the MDS that are administered to provide information comparable to items asked during the Community Questionnaire. Examples include items about prostate exams and mammograms, Instrumental Activities of Daily Living, vaccinations, smoking history, and general health. Interviewers ask these questions of someone knowledgeable about the beneficiary's care or find the information in the medical chart.

Since the 2019 Facility Instrument redesign and the usage of administrative data for Medicare and/or Medicaid certified facilities, only facilities that do not report a CCN receive the full HS section.

Operational and Procedural

The **Interviewer Remarks (IR)** section captures additional metadata about the interview, as recorded by the interviewer. This includes comments the interviewer may have about the interviewing situation and notes to themselves for use in gaining cooperation in the future. Data from this section are not included in the Survey File or the Cost Supplement File.

Missing Data Sections

There are three additional sections, called missing data sections, which are activated when essential survey information is coded as "don't know" or "refused" in the Facility Questionnaire (FQ), Residence History (RH), or Background (BQ) sections. The missing data sections prompt the interviewer for the specific piece of information that is missing. There are no new questions in the missing data sections, just repeats of questions initially asked in the FQ, RH, or BQ. Examples of the type of missing information that activate the missing data sections are the name of the facility or date of death.

The purpose of the missing data sections is to reduce item non-response for key variables in a modular, flexible format. If the interviewer is able to obtain the missing information from another Facility staff member or from a different medical document, then the interviewer uses the missing data section to capture a non-missing response for the key questionnaire item without modifying responses for the other already-completed items in the FQ, RH, and BQ sections. If the interviewer is unable to obtain the missing information, either "don't know" or "refused" is entered in the missing data sections.

¹¹ "Long-Term Care Facility Resident Assessment Instrument 3.0 User's Manual v.1.17.1," Centers for Medicare & Medicaid Services, October 2019.

¹² <https://www.cms.gov/data-research/computer-data-systems/minimum-data-sets-3.0-public-reports>

The missing data sections are:

- Facility Questionnaire Missing Data (FQ_MD): collects data missing from the FQ section of the interview;
- Residence History Questionnaire Missing Data (RH_MD): collects data missing from the RH section;
- Background Questionnaire Missing Data (BQ_MD): collects data missing from the BQ section.

5.3.4 Facility Topical Questionnaire Sections

In addition to the Core content, there are two Topical questionnaire sections that capture data on COVID-19 topics at the facility- and beneficiary-level. Each Topical section is described below, organized by information collected. Exhibit 5.3.4 lists the Topical sections and administration schedule.

Exhibit 5.3.4: 2022 Data Year MCBS Facility Topical Sections by Data File and Data Collection Schedule

Section Group	Abbr.	Section Name	LDS [§]	Data Collection Schedule
COVID-19	CV	COVID-19 Beneficiary	SF	All Seasons
	FC*	COVID-19 Facility-Level	SF	Fall 2022*

*The FC section is also administered to Community-to-Facility and Facility-to-Facility crossover cases each season.

[§]Limited Data Set (LDS) indicates the file where the questionnaire data appears (i.e., SF = Survey File, CS = Cost Supplement File).

COVID-19

The COVID-19 Supplement Topical sections capture key characteristics on the impact of the COVID-19 pandemic on long-term care facilities and Medicare beneficiaries.

The **COVID-19 Facility-Level (FC)** section collects information on topics that assess key ways in which COVID-19 has impacted facilities that serve Medicare beneficiaries. The FC section is separated into three main topics: availability of current telehealth services inside and outside of the facility, facility measures to prevent and control the spread of COVID-19, and mental health and social and recreational services offered inside and outside of the facility.

The FC section is completed for each beneficiary in the facility regardless of whether multiple beneficiaries live in the same facility. The Facility administrator is typically the most knowledgeable respondent for this section, although other Facility respondents may include (but are not limited to) social workers or MDS coordinators.

The **COVID-19 Beneficiary (CV)** section collects information on topics related to the beneficiary's utilization of COVID-19 testing, COVID-19 medical care, and COVID-19 vaccine utilization. The CV section is completed for each alive beneficiary in the facility.

The best Facility respondent for this section is someone directly involved with the beneficiary's care or someone who is familiar with their medical records.

6. SAMPLING

6.1 Medicare Population Covered by the 2022 MCBS Data

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 from both populations are residents of foreign countries and U.S. possessions and territories.

The beneficiaries included in the 2022 MCBS LDS releases represent a random cross-section of all beneficiaries who were ever enrolled in either Part A or Part B of the Medicare program for any portion of 2022. A subset of these beneficiaries represents a random cross-section of all beneficiaries who were continuously enrolled from January 1, 2022 up to and including interviews conducted during Fall 2022. The ever enrolled and continuously enrolled populations are described in further detail below:

- 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 2022 as well as beneficiaries who dis-enrolled or died prior to their fall interview.¹³ The ever enrolled population includes beneficiaries who were enrolled in Medicare for at least one day at any point during 2022.
- The continuously enrolled population represents only individuals continuously enrolled in Medicare from January 1, 2022 up to and including their fall interview; this specifically excludes beneficiaries who enrolled during the calendar year 2022 and beneficiaries 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 Survey File and Cost Supplement File represent four separate MCBS panels identified by the year in which the panel was selected and first interviewed (i.e., for 2022 LDS files, the 2019, 2020, 2021, and 2022 Panels). Exhibit 6.1 shows the composition of each of the four panels included in the 2022 data files.

Exhibit 6.1: 2022 MCBS Composition of Panels Contributing to the LDS Data Files

Data Year (Fall)	Number of Beneficiaries Selected
2019	11,615
2020	15,952
2021	15,950
2022	17,139

¹³ Note that data collection for beneficiaries who enrolled during 2023 and died in 2023 after enrollment but before their fall interview was still pursued through attempts at conducting proxy interviews.

Exhibit 6.1.1 presents the aggregated estimates of the size of the two Medicare populations overall and by sex and race.¹⁴ Exhibits 6.1.2 and 6.1.3 present estimates of the size of the continuously enrolled and ever enrolled Medicare populations by race, and age (as of December 31, 2022) for male and female beneficiaries.

Exhibit 6.1.1: 2022 Total Estimated Number of Medicare Beneficiaries by Sex and Race*

Group	Subgroup	Continuously Enrolled	Ever Enrolled
Overall Total		60,670,061	66,117,020
Sex	Male Total	27,546,896	30,119,259
	Female Total	33,123,166	35,997,761
Race	White non-Hispanic Total	44,193,670	46,686,023
	Black non-Hispanic Total	6,530,601	6,731,537
	Hispanic Total	4,810,682	5,031,208
	Other Total [†]	5,135,108	7,668,252

SOURCE: Beneficiary race/ethnicity were sourced from administrative data in the Sample Control File and the weights were sourced from the 2022 Survey File.

* 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.

¹⁴ 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.

Exhibit 6.1.2: 2022 Estimated Number of Male Medicare Beneficiaries by Race and Age*

Race	Age as of 12/31/2022	Continuously Enrolled	Ever Enrolled
White non-Hispanic	<45	463,628	498,110
	45-64	1,829,168	1,874,216
	65-69	4,566,586	4,984,169
	70-74	5,161,794	5,234,034
	75-79	3,960,879	3,995,280
	80-84	2,403,344	2,611,935
	85+	1,837,776	2,149,103
Black non-Hispanic	<45	157,033	157,743
	45-64	599,626	584,612
	65-69	655,810	722,798
	70-74	642,484	647,826
	75-79	395,880	394,080
	80-84	195,674	210,090
	85+	131,649	173,889
Hispanic	<45	77,686	84,885
	45-64	299,460	309,338
	65-69	516,703	542,381
	70-74	536,124	533,665
	75-79	328,315	320,240
	80-84	176,174	194,826
	85+	126,262	164,005
Other[†]	<45	122,965	179,115
	45-64	201,007	377,499
	65-69	957,495	1,932,536
	70-74	602,980	621,345
	75-79	402,849	405,790
	80-84	96,740	99,484
	85+	100,803	116,264

SOURCE: Beneficiary age and race/ethnicity were sourced from administrative data in the Sample Control File and the weights were sourced from the 2022 Survey File.

* 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.

Exhibit 6.1.3: 2022 Estimated Number of Female Medicare Beneficiaries by Race and Age*

Race	Age as of 12/31/2022	Continuously Enrolled	Ever Enrolled
White non-Hispanic	<45	337,232	364,752
	45-64	1,664,762	1,698,215
	65-69	5,068,442	5,636,359
	70-74	6,041,338	6,226,828
	75-79	4,911,027	4,856,714
	80-84	3,015,662	3,169,584
	85+	2,932,032	3,386,724
Black non-Hispanic	<45	134,662	140,322
	45-64	608,332	606,675
	65-69	895,281	910,921
	70-74	883,401	883,823
	75-79	567,053	555,614
	80-84	322,982	326,587
	85+	340,734	416,557
Hispanic	<45	86,361	85,937
	45-64	316,950	323,965
	65-69	691,422	768,166
	70-74	616,106	633,517
	75-79	429,357	430,015
	80-84	283,630	294,407
	85+	326,131	345,861
Other[†]	<45	117,863	140,836
	45-64	225,783	392,779
	65-69	1,091,911	2,143,185
	70-74	495,404	518,909
	75-79	344,272	338,905
	80-84	174,502	184,610
	85+	200,531	216,994

SOURCE: Beneficiary age and race/ethnicity were sourced from administrative data in the Sample Control File and the weights were sourced from the 2022 Survey File.

* 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.

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 was in one of the 48 contiguous states (excludes Alaska and Hawaii) or the District of Columbia. For example, for Fall Rounds 2019, 2020, 2021, and 2022 (the four rounds in which the 2019, 2020, 2021, and 2022 Panels, included in the 2022 MCBS data, were selected), the targeted population included individuals enrolled as of December 31 of 2019, 2020, 2021, and 2022 respectively.

The universe of beneficiaries for the MCBS is divided into seven sampling strata based on age as of December 31 of the sampling year in order to include all beneficiaries enrolling during the sampling year. The age categories are under 45, 45 to 64, 65 to 69, 70 to 74, 75 to 79, 80 to 84, and 85 or older. The strata also separate Hispanic and non-Hispanic beneficiaries by age group. The 14 strata in 2022 are depicted in Exhibit 6.2.1.¹⁵

Exhibit 6.2.1: 2022 MCBS Sampling Strata

Hispanic	Non-Hispanic
Under 45 years Hispanic	Under 45 years non-Hispanic
45 - 64 Hispanic	45 - 64 non-Hispanic
65 - 69 Hispanic	65 - 69 non-Hispanic
70 - 74 Hispanic	70 - 74 non-Hispanic
75 - 79 Hispanic	75 - 79 non-Hispanic
80 - 84 Hispanic	80 - 84 non-Hispanic
85 and over Hispanic	85 and over non-Hispanic

Additionally, in the 2019, 2020, 2021, and 2022 Panels, beneficiaries residing within the U.S. who were Hispanic (based on a Hispanic ethnicity classification code in the Medicare enrollment data; see Eicheldinger¹⁶ for more details) were oversampled to improve precision of estimates for this group.¹⁷ See the *MCBS Methodology Report* for more information about this oversample. Exhibit 6.2.2 displays the beneficiaries selected as part of the 2022 Panel, by age and ethnicity.

¹⁵ Note that the MCBS surveys beneficiaries living in community (e.g., households) and in facility (e.g., nursing home) settings; however, residence status is not known at the time of sampling and is therefore not included among the MCBS sampling strata.

¹⁶ 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.

¹⁷ 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/data-research/research/medicare-current-beneficiary-survey/data-documentation-codebooks>.

Exhibit 6.2.2: 2022 Panel of Selected Beneficiaries by Hispanic and Non-Hispanic Ethnicity Classification and Age Category*

Age Category as of 12/31/2022	TOTAL Sample Size	TOTAL Weighted	Hispanic Sample Size	Hispanic Weighted	Non-Hispanic Sample Size	Non-Hispanic Weighted
Under 45 years	1,216	1,668,565	124	174,013	1,092	1,494,552
45-64 years	1,517	6,239,898	196	551,759	1,321	5,688,139
65-69 years	3,099	17,444,495	355	1,393,024	2,744	16,051,471
70-74 years	2,536	15,885,424	296	1,238,731	2,240	14,646,693
75-79 years	2,656	11,515,452	309	795,446	2,347	10,720,005
80-84 years	2,942	7,207,845	352	560,846	2,590	6,646,999
85+ years	3,173	7,044,936	363	482,823	2,810	6,562,113
Total	17,139	67,006,615	1,995	5,196,643	15,144	61,809,972

SOURCE: Beneficiary age and race/ethnicity were sourced from administrative data in the Sample Control File and the weights were sourced from the 2022 Survey File.

6.2.1 Eligibility: Medicare Population Covered by the 2022 LDS

Beneficiaries who became eligible for Medicare Part A or B and enrolled anytime during the year were eligible to be sampled as part of the annual panel.¹⁸ Thus, the 2022 Survey File and Cost Supplement File includes data from the 2019, 2020, 2021, and 2022 Panels. The inclusion of the current-year enrollees allows data to be released in a timelier manner; the Survey File LDS is released 12-15 months after the end of data collection, and the Cost Supplement File LDS is released 15-18 months after the end of data collection.

6.3 Three-Stage Cluster Design

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 MCBS sample is annually "supplemented" during the fall round to account for attrition (deaths, dis-enrollments, refusals) and newly enrolled persons. Each annual supplement is referred to as the Incoming Panel sample. For more information about the selection of the PSUs and SSUs, see the *MCBS Methodology Report*.

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.¹⁹ For each continuing beneficiary, the survey questions corresponding to the Survey File data release are administered in the fall of the data collection year. Similarly, for beneficiaries new to the MCBS, the survey questions are administered as part of the initial fall Baseline interview. Exhibit 6.4 provides a brief summary of the number of selected beneficiaries and the inclusion criteria for the 2019 through 2022 Panels.

¹⁸ These beneficiaries are referred to as "current-year enrollees."

¹⁹ The MCBS 2022 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*.

Exhibit 6.4: 2022 MCBS Sample Selection for the LDS Releases

Panel	# of Selected Beneficiaries	Previously Enrolled Beneficiaries Still Alive as of January 1 of Panel Year	Current-Year Enrollees
2019	11,615	Enrolled before 1/1/2019	Enrolled 1/1/2019 – 12/31/2019
2020	15,952	Enrolled before 1/1/2020	Enrolled 1/1/2020 – 12/31/2020
2021	15,950	Enrolled before 1/1/2021	Enrolled 1/1/2021 – 12/31/2021
2022	17,139	Enrolled before 1/1/2022	Enrolled 1/1/2022 – 12/31/2022

SOURCE: 2022 MCBS Internal Sample Control File

7. DATA PRODUCTS & DOCUMENTATION

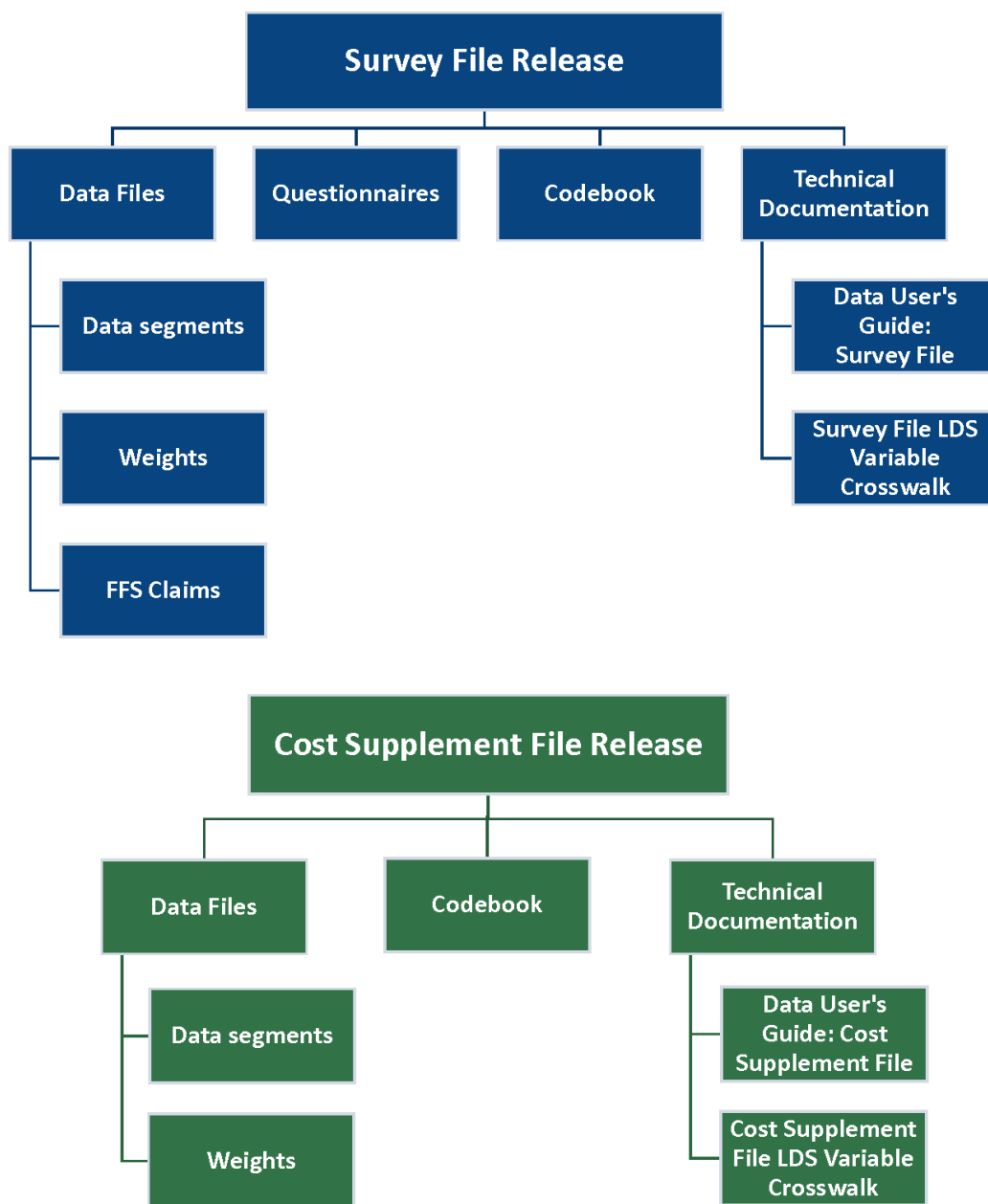
7.1 Contents of Data Release

MCBS data are made available via releases of annual files. For each data year, two annual LDS releases (the Survey File and the Cost Supplement File) and two PUFs (based on the Survey File and Cost Supplement File, respectively) are planned.²⁰ The LDS releases contain multiple files, called segments, which are easily linkable through a common beneficiary key ID. The 2022 Survey File LDS contains over 6,000 variables across 47 segments and the 2022 Cost Supplement File LDS contains over 600 variables across 14 segments.

Detailed descriptions of each segment, including the core contents of each segment, key variable definitions, and special notes on new variables, recodes, and administrative sources for select variables can be found in this document (see Section 10) and corresponding information for the Cost Supplement File can be found in the *MCBS Data User's Guide: Cost Supplement File*.

Exhibit 7.1 displays the components of each LDS release. Both the Survey File and Cost Supplement File contain data segments, codebooks, questionnaires, and technical documentation. The Survey File release contains the FFS claims data, which provide CMS administrative information on medical services and payments paid by Medicare claims; PDE events for Medicare Part D are not included and claims data for Medicare Advantage beneficiaries are not available. While users can conduct analyses with the Survey File alone, users interested in the Cost Supplement File data will need both LDS files to link cost and utilization variables with demographic or health insurance coverage variables.

²⁰ In addition to the annual MCBS Survey File PUF and MCBS Cost Supplement File PUF, CMS has released three special topic PUFs with data from the three MCBS COVID-19 Community Supplements, which correspond to the 2019 and 2020 data years. For more information, see the Data User's Guides available here: <https://www.cms.gov/data-research/statistics-trends-and-reports/mcbs-public-use-file>.

Exhibit 7.1: 2022 Contents of Data Releases**7.1.1 2022 MCBS Survey File**

The Survey File contains data collected directly from respondents and supplemented by administrative items plus the facility (non-cost) information and FFS claims. The Survey File includes multiple topic-related segments, including health status and limitations, access to care, health insurance coverage, and household characteristics. The Survey File also includes information on Facility interviews, including a residence timeline, facility characteristics, and assessment (Minimum Data Set) measures. Finally, Topical questionnaire sections (e.g., beneficiary knowledge, drug coverage) are included with this release. To facilitate analysis, the information collected in the survey is augmented with data on the use and program cost of Medicare services from Medicare claims data and administrative data. The Survey File includes beneficiaries enrolled for at least

one day in 2022 who completed an interview in 2022 or Winter 2023, or who died during 2022. Beneficiaries who refused to complete a later interview or became nonrespondents during 2022 are excluded.

Exhibit 7.1.1 displays each segment included in the Survey File including the **segment abbreviation**, **brief description**, and **information on weights or other special notes**.

The **Data Source** column describes the source of the data on the segment. The three possible sources are the Community Questionnaire (CQ), Facility Instrument (FI), and Administrative Records (AR). Each LDS segment can have any combination of these sources. Data source reflects where the data came from, not where the beneficiary was living. For example, a beneficiary could have lived in both settings during the year, but the data for that beneficiary available on the ACCESSCR segment came from their Community interview only.

The **Quex Section** column lists the specific questionnaire sources for the LDS segment. Please note that not all variables from the questionnaire are released on the segments. Some questionnaire items are combined or recoded to create the LDS variable. Data users will see these derived variables noted in the codebooks preceded with the character "D", such as D_ERVIST.

Season indicates the round (winter, summer, fall) and year when the questionnaire was administered.

Panel describes whether the questionnaire sections that provide the data for each segment are fielded for baseline respondents (Base.), continuing respondents (Cont.), or all panels (All). If the segment consists of administrative CMS data, then the cell indicates all panels are included.

Unit of Observation indicates what each row in the segment represents. For example, the ASSIST segment provides multiple rows per BASEID for each person reported as helping the beneficiary in the data year.

A list of equivalent historic segments from the 1991-2013 data release structure is provided in Appendix E.

Exhibit 7.1.1: 2022 MCBS Survey File Segments and Contents

Survey File Segment (Abbrev)	Description	Data collection and special weights notes	Data Source*	Quex Section	Season	Panel*	Unit of Observation
Access to Care (ACCESSCR)	Information on ability to obtain health care, delay of care related to costs, and reasons for not obtaining needed health care.		CQ	HFQ	Fall	All	Beneficiary
Access to Care Medical Appointments (ACCSSMED)	Information on medical and dental visit experiences and forgone medical, dental, vision, hearing, and mental health care and prescription medicines.	The data collected in Winter 2023 are released with the 2022 Survey File given that the reference period is 2022. Special non-response adjustment weights are included with this file.	CQ	ACQ, DVH, MPQ, PMQ	Winter (ACQ) ³ All (DVH, MPQ, PMQ)	Cont.	Beneficiary
Administrative Utilization Summary (ADMNUTLS)	Summarized administrative information on Medicare, program expenditures, and utilization.		AR	n/a (Admin data)	n/a	All	Beneficiary
Assistance (ASSIST)	Information on the person helping and type of assistance that the beneficiary may receive performing ADLs and IADLs.		CQ	ENS, HFQ	All (ENS) Fall (HFQ)	All	Helper by beneficiary
Chronic Conditions (CHRNCOND)	Information on chronic and other diagnosed medical conditions.		CQ	HFQ, PVQ	Fall (HFQ, PVQ) ³	All	Beneficiary
Chronic Conditions Flags (CHRNCDL)	FFS Chronic Condition Flag Records and FFS Chronic and other Disabling Flag records from administrative data sources.		AR	n/a (Admin data)	n/a	n/a	Beneficiary
Chronic Pain (CHRNPAIN)	Information on experiences with chronic pain and non-medication related chronic pain management techniques.	The data collected in Summer 2023 are released with the 2022 Survey File given that the reference period is 2022. Special non-response adjustment weights are included with this file.	CQ	CPQ	Summer	Cont.	Beneficiary
Cognitive Measures (COGNFUNC)	Measures of cognitive functioning.		CQ	CMQ	Fall	All	Beneficiary

Survey File Segment (Abbrev)	Description	Data collection and special weights notes	Data Source*	Quex Section	Season	Panel*	Unit of Observation
COVID-19 Topical (COVIDTOP)	Information on COVID-19 vaccination, testing, diagnosis, symptoms, and prevention.	The data collected in Winter and Summer 2023 are released with the 2022 Survey File given that the reference period is 2022. Special non-response adjustment weights are included with this file.	CQ	CVQ	Winter, Summer	All	Beneficiary
COVID-19 Experience (COVIDEXP)	Information on COVID-19 vaccination, testing, diagnosis, symptoms, and prevention.		CQ	CVQ	Fall	All	Beneficiary
Demographics (DEMO)	Demographic information.		CQ, FI, AR	ENS, DIQ, INQ, BQ, RH	All (ENS, INQ, RH) Fall ¹ (DIQ, BQ)	All (ENS, INQ, BQ, RH) Base. (DIQ)	Beneficiary
Diabetes (DIABETES)	Information on diabetes management such as insulin usage.		CQ	HFQ	Fall	All	Beneficiary
Facility Assessments (FACASMNT)	Assessment information conducted while the beneficiary was living in a Medicare approved or non-Medicare approved facility.		FI, AR	HS	Fall ²	All	Beneficiary
Facility Characteristics (FACCHAR)	Primarily information from the Facility Questionnaire with Skilled Nursing Facility (SNF) stay information for beneficiaries living in the community and in facilities incorporated.		FI, AR	BQ, FQ, RH	Fall ¹ (BQ) All (FQ, RH)	All	Facility by beneficiary
Falls (FALLS)	Information on injuries and attitudes about falls.		CQ	HFQ	Fall	All	Beneficiary
Food Insecurity (FOODINS)	Information on access to sufficient food.	The data collected in Summer 2023 are released with the 2022 Survey File given that the reference period is 2022. Special non-response adjustment weights are included with this file.	CQ	IAQ	Summer	Cont.	Beneficiary

Survey File Segment (Abbrev)	Description	Data collection and special weights notes	Data Source*	Quex Section	Season	Panel*	Unit of Observation
General Health (GENHLTH)	Information on general health status and functioning such as height and weight.		CQ	HFQ	Fall	All	Beneficiary
Health Insurance Summary (HISUMRY)	Administrative information on the characteristics of insurance coverage.		CQ, AR	HIQ	All	All	Beneficiary
Health Insurance Timeline (HITLINE)	Information on insurance plans and the coverage eligibility timeline as well as information regarding premiums and covered services.		CQ, FI, AR	CPS, ENS, HIQ, NSQ, STQ, IN	All (CPS, HIQ, NSQ, STQ) Fall ² (IN)	Cont. (CPS, NSQ, STQ) Both (HIQ)	Plan type by beneficiary
Household Characteristics (HHCHAR)	Information on household composition and home.		CQ	ENS, HAQ	All (ENS) Fall (HAQ)	All	Beneficiary
Income and Assets (INCASSET)	Information on income and assets.	The data collected in Summer 2023 are released with the 2022 Survey File given that the reference period is 2022. Special non-response adjustment weights are included with this file.	CQ	IAQ	Summer ³	Cont.	Beneficiary
Interview Characteristics (INTERV)	Information on interview characteristics.		CQ, FI	END, ENS, INQ, IRQ	All	All	Interview by beneficiary
MA Plan Questions (MAPLANQX)	Information on access to and satisfaction with care for beneficiaries enrolled in Medicare Part C.		CQ	HIQ	All	All	Beneficiary
Medicare Plan Beneficiary Knowledge (MCREPLNQ)	Information on experiences with the Medicare open enrollment period and knowledge about Medicare-covered expenses.	The data collected in Winter 2023 are released with the 2022 Survey File given that the reference period is 2022. Special non-response adjustment weights are included with this file.	CQ	KNQ	Winter ³	Cont.	Beneficiary
Minimum Data Set (MDS3)	Assessment information conducted while the beneficiary was living in an approved Medicare facility.		AR	n/a (Admin data)	n/a	n/a	Assessment by beneficiary

Survey File Segment (Abbrev)	Description	Data collection and special weights notes	Data Source*	Quex Section	Season	Panel*	Unit of Observation
Mental Health (MENTHLTH)	Information on mental health such as feelings of anxiety or depression.		CQ	HFQ	Fall	All	Beneficiary
Mobility (MOBILITY)	Information on the use of available transportation options and whether health status affects their daily travel.		CQ	MBQ	Fall	All	Beneficiary
Multiple Year Enrollment (MYENROLL)	Up to five years of beneficiary enrollment information with monthly flags related to Part D and LIS enrollment, dual eligibility status, and type of Medicare coverage.		AR	n/a (Admin data)	n/a	All	Beneficiary
Nagi Disability (NAGIDIS)	Information on difficulties with performance of activities of daily living.		CQ	HFQ	Fall	All	Beneficiary
Nicotine and Alcohol (NICOALCO)	Information on the prevalence and frequency of alcohol and nicotine use.		CQ	NAQ	Fall	All	Beneficiary
Outcome and Assessment Information (OASIS)	Assessment information conducted while the beneficiary was receiving home health services.		AR	n/a (Admin data)	n/a	n/a	Assessment by beneficiary
Patient Activation (PNTACT)	Information on the degree to which beneficiaries actively participate in their health care and decisions concerning care.	Special non-response adjustment weights are included with this file.	CQ	SCQ	Fall	All	Beneficiary
Preventive Care (PREVCARE)	Information on preventive services such as vaccinations and routine screening procedures.		CQ	HFQ, PVQ	Fall (HFQ) All (PVQ) ³	All	Beneficiary
Residence Timeline (RESTMLN)	Information on where the beneficiary lived over the course of the year.		CQ, FI	HHQ, IPQ, IUQ	All	Cont.	Beneficiary

Survey File Segment (Abbrev)	Description	Data collection and special weights notes	Data Source*	Quex Section	Season	Panel*	Unit of Observation
RX Medications (RXMED)	Information on prescription medication access and satisfaction with and knowledge about Medicare Part D.	The data collected in Summer 2023 are released with the 2022 Survey File given that the reference period is 2022. Special non-response adjustment weights are included with this file.	CQ	RXQ	Summer ³	Cont.	Beneficiary
Satisfaction with Care (SATWCARE)	Information on satisfaction with different aspects of health care.		CQ	SCQ	Fall	Cont. (MPQ, PMQ) Both (SCQ)	Beneficiary
Social Determinants of Health (SDOH)	Information on beneficiaries' residential history and social determinants of health		AR	n/a (Admin data, external sources) ⁴	n/a	All	Beneficiary
Telemedicine (TELEMED)	Information on telemedicine visit availability and usage.	The data collected in Winter 2023 are released with the 2022 Survey File given that the reference period is 2022. Special non-response adjustment weights are included with this file.	CQ	TLQ	Winter ³	Cont.	Beneficiary
Usual Source of Care (USCARE)	Information on where and how the beneficiary typically seeks medical care.	The data collected in Winter 2023 are released with the 2022 Survey File given that the reference period is 2022. Special non-response adjustment weights are included with this file.	CQ	USQ	Winter ³	Cont.	Beneficiary
Vision and Hearing (VISHEAR)	Information on eye health and hearing status.		CQ	HFQ	Fall	All	Beneficiary
Weights (CENWGTS) (EVRWGTS) (LNG2WGTS) (LNG3WGTS) (LNG4WGTS)	The weights segments include longitudinal weights for the continuously enrolled population, general-purpose cross-sectional weights, a series of replicate weights, and weights to represent the ever enrolled population.		CQ, FI	n/a	n/a	All	Beneficiary

Survey File Segment (Abbrev)	Description	Data collection and special weights notes	Data Source*	Quex Section	Season	Panel*	Unit of Observation
COVID-19 Facility Beneficiary-Level (FBENCVFL)	Information on COVID-19 diagnosis, testing, and care received by beneficiaries living in a facility during the fall of 2022 and winter of 2023.		FI	CV	Fall, Winter	All	Beneficiary
COVID-19 Facility Facility-Level (FFACCVFL)	Information on ways COVID-19 impacted facilities that serve Medicare beneficiaries during the fall of 2022 and winter of 2023.		FI	FC	Fall, Winter ²	All	Facility by beneficiary
Fee-for-Service Claims (FFS)	Abbreviated FFS claims data. Additional claims-like data will be included as they become available in subsequent years (e.g., Encounter Data, Medicaid claims data).		AR	n/a (Admin data)	n/a	All	Beneficiary

* = Data source describes the source of the data on the segment. The three possible sources are the Community Questionnaire (CQ), Facility Instrument (FI), and Administrative Records (AR). Each LDS segment can have any combination of these sources. Data source reflects where the data came from, not where the beneficiary was living. For example, a beneficiary could have lived in both settings during the year but the data for that beneficiary available on the ACCESSCR segment came from their Community interview only.

** = Panel describes whether the questionnaire sections that provide the data for each segment are fielded for baseline respondents, continuing respondents, or both.

1. The BQ section is also administered to Community-to-Facility Crossover cases each season.
2. The FC, IN, and HS sections are also administered each season to Community-to-Facility cases, Facility-to-Facility cases, and for beneficiaries living in a Facility whose last interview was a Community interview and who completed a Facility interview in a prior round cases.
3. These sections are administered in rounds following the current data year given that the reference period is the prior year and data are included in the prior year data files.
4. The SDOH segment includes data sourced from the Social Deprivation Index, CDC/ATSDR Social Vulnerability Index, and AHRQ's Social Determinants of Health Database.

7.1.2 2022 MCBS Cost Supplement File

The Cost Supplement File contains both individual event and summary files and can be linked to the Survey File to conduct analyses on health care cost and utilization. The Cost Supplement File links survey-reported events to Medicare FFS claims and provides a comprehensive picture of health services received, amounts paid, and sources of payment, including those not covered by Medicare. Survey-reported data include information on the use and cost of all types of medical services, as well as information on supplementary health insurance costs. Medicare FFS claims data include administrative and billing information on the use and cost of inpatient hospitalizations, outpatient hospital care, physician services, home health care, durable medical equipment, skilled nursing home services, hospice care, and other medical services.²¹ The Cost Supplement File can support a broader range of research and policy analyses on the Medicare population than would be possible using either survey data or administrative claims data alone.

The Cost Supplement File contains a subset of the beneficiaries included in the Survey File who have complete cost and utilization data for the year. For the 2022 MCBS data year, the Cost Supplement File includes beneficiaries sampled in the 2019 through 2021 Panels, plus members of the 2022 Panel who were enrolled in Medicare during 2022 for at least one day.

For beneficiaries enrolled in Medicare Advantage, cost and utilization information is available. As is done with services not covered by Medicare (e.g., most dental, vision, and hearing care), when a beneficiary reports health care events, the MCBS uses the explanation of benefits (EOB) form from Medicare Advantage providers to report the payments, as well as the capitation information from the administrative data for total Medicare Advantage Payments. Actual claims-based information for MA beneficiaries, referred to as encounter data, is not currently available for these individual events. The Cost Supplement File undergoes a careful reconciliation process to separately identify and flag health care services reported: 1) from the survey alone, 2) from the claims data alone, and 3) from both sources. This process results in a file with a much more complete and accurate picture of health services received, amounts paid, and sources of payment. Due to the added processing time required to reconcile survey reported events with the claims data, this file is generally released 18 months after the close of the calendar year for data collection.

Exhibit 7.1.2 displays each segment included in the Cost Supplement File along with the abbreviation, description, and the equivalent historic segment from the 1991-2013 data release structure.

²¹ Only Medicare claims for beneficiaries enrolled in Medicare Fee-for-Service (FFS, often called “traditional” Medicare), are available for linkage; similar claims information for Medicare Advantage (MA) beneficiaries is not available. To the extent that health care use and costs may be underreported in the survey or reported differentially between FFS and MA beneficiaries, this will be reflected in the data as MA beneficiaries’ information will not be supplemented by claims data.

Exhibit 7.1.2: 2022 MCBS Cost Supplement File Segments and Contents

Cost Supplement Segment (Abbrev)	Description	Data collection and special weights notes	Data Source*	Quex Section	Season	Panel**	Unit of Observation
Dental Utilization Events (DUE)	Contains individual dental events reported during a Community interview or created from Medicare claims data.		CQ, AR	DVH, US	All	All	One record per beneficiary per event (defined as a single visit to the dentist)
Facility Events (FAE)	Contains individual facility events reported during a Facility interview.	There is one record for each stay that occurred at least partly in the data year.	FI, AR	RH, US, EX	All	All	One record per beneficiary per stay in a long-term care facility
Hearing Utilization Events (HUE)	Contains individual hearing care events reported during a Community interview or created from Medicare claims data.		CQ, AR	DVH, US	All	All	One record per beneficiary per event (defined as a single visit to a hearing care provider)
Inpatient Hospital Events (IPE)	Contains individual inpatient hospital events reported during a Community interview or created from Medicare claims data.		CQ, AR	IUQ, IPQ, ERQ, OPQ, US	All	All	One record per beneficiary per admission
Institutional Events (IUE)	Contains individual short-term facility (usually skilled nursing facility) stays reported during a Community interview or created from Medicare claims data.		CQ, AR	IUQ, IPQ, US	All	All	One record per beneficiary per admission
Medical Provider Events (MPE)	Contains individual events for a variety of medical services, equipment, and supplies reported during a Community interview or created from Medicare claims data.		CQ, AR	ERQ, IPQ, MPQ, OMQ, OPQ, US	All	All	One record per beneficiary per event (defined as a separate visit, procedure, service, or a supplied item for a survey-reported event)
Outpatient Hospital Events (OPE)	Contains individual outpatient hospital events reported during a Community interview or created from Medicare claims data.		CQ, AR	OPQ	All	All	One record per beneficiary per event (defined as a single outpatient visit)

Cost Supplement Segment (Abbrev)	Description	Data collection and special weights notes	Data Source*	Quex Section	Season	Panel**	Unit of Observation
Prescribed Medicine Events (PME)	Contains individual outpatient prescribed medicine events reported during a Community interview or created from Medicare claims data.		CQ, AR	PMQ, DVH, ERQ, IPQ, OPQ, MPQ	All	All	One record per beneficiary per prescribed medicine (defined as a single prescribed medicine)
Vision Utilization Events (VUE)	Contains individual vision care events reported during a Community interview or created from Medicare claims data.		CQ, AR	DVH, US	All	All	One record per beneficiary per event (defined as a single visit to a vision care provider)
Person Summary (PS)	Summarization of utilization and expenditures by type of service and summarization of expenditures by payer, yielding one record per person.		CQ, FI, AR	all utilization including HHQ, US	All	All	One record per beneficiary
Service Summary (SS)	Summarization of the nine individual event files along with one record for home health and one record for hospice utilization, yielding a total of 11 summary records per person.		CQ, FI, AR	all utilization including HHQ, US	All	All	11 records per beneficiary
CSEVWGTS	Contains cross-sectional full-sample and replicate weights representing the 2022 ever enrolled population.		CQ/FI	N/A	All	All	One record per beneficiary
CSL2WGTS CSL3WGTS	Contains longitudinal full-sample and replicate weights for the multi-year ever enrolled population. The CSL2WGTS file includes the two-year longitudinal weights for the population ever enrolled at any time during both 2021 and 2022. The CSL3WGTS file includes the three-year longitudinal weights for the population ever enrolled at any time during 2020, 2021, and 2022.		CQ/FI	N/A	All	All	One record per beneficiary

* = Data source describes the source of the data on the segment. The three possible sources are the Community Questionnaire (CQ), Facility Instrument (FI), and Administrative Records (AR). Each LDS segment can have any combination of these sources.

** = Panel describes whether the questionnaire sections that provide the data for each segment are fielded for baseline respondents, continuing respondents, or both.

The Cost Supplement segments are assembled at three levels:

- The Event level reports all payers, costs, and utilization at the most detailed level available (one observation per event per person).
- The Service Summary level summarizes all payers, costs, and utilization for a person at the service level (one observation per service type per person).
- The Person Summary level summarizes all payers and costs across service categories and summarizes type of service amounts (one observation per person).

The tri-level structure allows researchers to fit the research problem they are addressing to the available file summary levels, and potentially avoid having to process all the detailed event records in the file when summaries may suffice. For example, an analysis of differences in total health spending per person between men and women could use the person summary level, and thereby avoid having to process the more numerous event level records. Similarly, an analysis of differences in use of Medicare hospital payments by race could use the type of service summary records. Event level records would be used for more detailed analyses, for example, average length of long-term facility stays or average reimbursements per prescription drug type. For a more complete discussion of the tri-level file structure, see the *MCBS Data User's Guide: Cost Supplement File* document.

7.1.3 Using the Data

The MCBS data releases are made available in two formats: SAS formatted files and comma delimited files for use with Stata® and R®. Directions and sample SAS code are given below to help users read the dataset into SAS (see Appendix B). Files with programming code to create formats and labels are provided for both SAS users and for use with comma delimited files.

7.1.4 Research Claims Files

The fixed-length claims (also known as the research claims) are abbreviated versions of the full claim record layout. Each claim type has a subset of variables selected for their relevancy to data analysis of that service. Additionally, institutional claim types have a corresponding revenue center file that links back to the claim-level data file through a unique claim identifier. See Section 8.3: Claims Files for more on the claims file specifications.

There is one observation per data record for all of the MCBS claims files except the Physician/Supplier Claims and Durable Medical Equipment (DME) Claims. Those claim types treat each line item as a separate observation with the claim-level detail repeating for each line item.

7.2 Which File Do I Need?

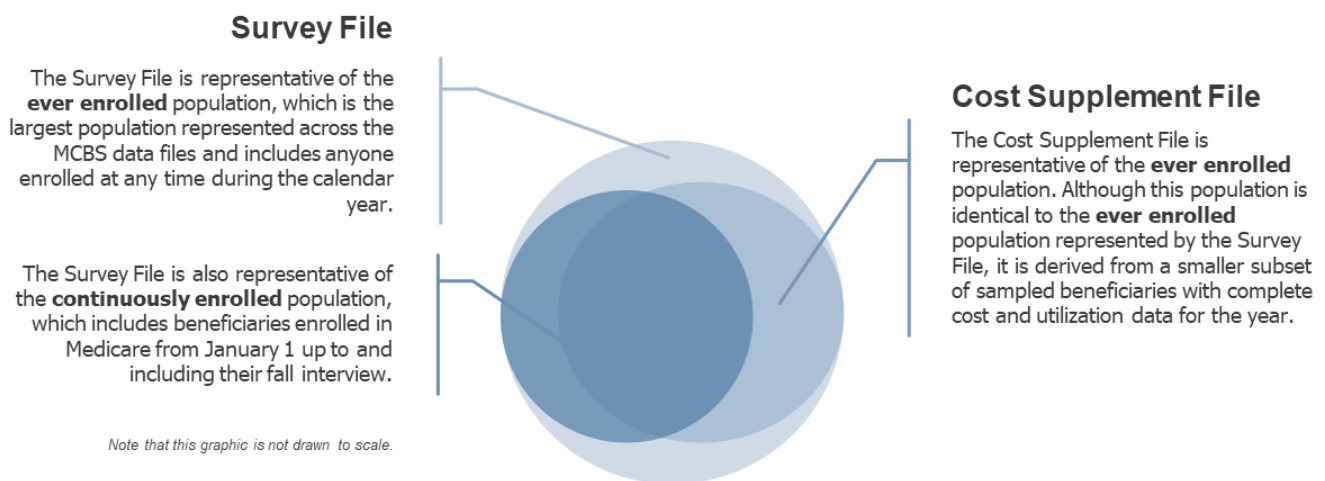
The identification of the target population for a given research question will influence both the selection of weights and the particular segments that a data user will need to conduct analyses. Exhibit 7.2 depicts the relationship between the beneficiaries included in the annual data releases.²² The ever enrolled population from the Survey File is the largest, including anyone enrolled at any time during the calendar year corresponding to the LDS data year. The continuously enrolled represent the population of beneficiaries who were enrolled continuously between January 1 through the completion of their fall interview. Beneficiaries who died during the year, newly enrolled beneficiaries who enrolled in Medicare during the year that they were sampled, and beneficiaries who have lost eligibility are not included in the continuously enrolled group. The

²² Exhibit 7.2 is not drawn to scale, but provided as a visual reference for the relationship of populations between data files.

ever enrolled represent the population of beneficiaries who were ever enrolled in Medicare for at least one day at any time during the year. The ever enrolled population includes beneficiaries who died or lost entitlement prior to completing the fall interview. Beneficiaries who first became enrolled in Medicare during the year are also included. Thus, the continuously enrolled beneficiaries are a subset of the ever enrolled beneficiaries. The Survey File LDS includes weight segments that allow for subsetting the data by the ever enrolled and continuously enrolled populations.

The Cost Supplement File is representative of the ever enrolled population, but is smaller than the Survey File population because it is derived from a smaller subset of sampled beneficiaries with complete cost and utilization data for the entire year. As does the Survey File, the Cost Supplement File includes a weight segment that allows for subsetting the data by the ever enrolled population.

Exhibit 7.2: MCBS Populations in Data Products



7.2.1 Survey File Only

Users who wish to focus on research questions around health-related topics, such as health status and access to care and/or Medicare FFS utilization, only need the Survey File.

7.2.2 Using Both Survey File and Cost Supplement File

To the extent that a data user needs demographic and health insurance information to conduct research on the cost and utilization of medical services, both the Survey File and the Cost Supplement File are required. Data users must also use the ever enrolled cost weights when analyzing any cost data from the Cost Supplement File combined with survey-reported information from the Survey File. For more information on using the weights, please see 9.4 Weighting.

7.2.3 Using Both Community and Facility Data

Analytic decisions about whether to include all beneficiaries regardless of residence status or beneficiaries living only in the community or only in facilities should be driven by both the research question and data limitations. However, as discussed in Sections 4 and 5, there are differences in the data collection protocols and questionnaire instruments for the MCBS Community and Facility components. Thus, caution should be observed when combining data across these populations to address questions requiring analysis of all Medicare beneficiaries.

In order to determine which population should be included in an analysis, the following steps are recommended:

1. Define the population based on the research question(s) and identify the living in community and living in facility populations. The variable INT_TYPE on the DEMO segment is the recommended variable for defining the two populations. See Section 10.3.11 for more information on INT_TYPE.
2. Identify the LDS segments and variables associated with each of the analysis' domains to determine what data are available for the Community and Facility components.
3. Assess whether the universe, level of measurement, and response categories for the variables of interest are similar for both Community and Facility components.
4. If needed, recode the LDS variables to align the coding between Community and Facility components and create analytic variables.
5. Merge the Community and Facility segments with the appropriate weights segments. Assess preliminary estimates for variation between community and facility.
6. Review MCBS documentation to determine if there are underlying differences in data collection and processing between community and facility that result in analytic limitations.
7. Conduct analysis and document any potential limitations.

For more information on using community and facility data, including a series of analytic examples with sample SAS code, see the *MCBS Advanced Tutorial on Using Community and Facility Data*. Data users can access this tutorial at <https://www.cms.gov/data-research/research/medicare-current-beneficiary-survey/data-briefs-tutorials>.

8. FILE STRUCTURE

8.1 LDS Specifications

The MCBS Survey File contains survey-collected data augmented with administrative and claims data to allow for analysis regarding the beneficiaries' health status, access to health care, satisfaction with health care, and usual source of care. The following information is represented in the MCBS Survey File: beneficiary demographics, household characteristics, access to care, satisfaction with care, usual source of care, health insurance timeline (shows types of insurances, the coverage eligibility, and what is covered), health status and functioning, and other Topical questionnaire sections like medical conditions and chronic pain, health behaviors, preventive services, interview characteristics, beneficiary knowledge of the Medicare program, residence timeline, facility characteristics, and income and assets.

In terms of Medicare eligibility and enrollment data, HITLINE provides monthly coverage indicators, coverage start and end dates, the type of plan, and the source of coverage information for the plan. HISUMRY also contains eligibility codes and detailed Medicare-Medicaid dual eligibility indicators.

8.2 File Structure

The Survey File segments can be divided into two subject matter groups: files containing survey data with related Medicare administrative variables and files containing Medicare FFS claims data. The claims records represent services provided during calendar year 2022 and processed by CMS. To facilitate analysis, the ADMNUTLS record contains a detailed summary of the utilization enumerated by these claims.

All MCBS segments begin with the same three variables: a unique number that identifies the person who was sampled (the BASEID), the survey reference year (in this release, a constant "2022"), and the version of release. These elements serve to identify the type of record and to provide a link to other types of records. To obtain complete survey information for an individual, a researcher must link together records for that individual from the various data files using the variable BASEID. Beneficiaries may not have a record on every data file. Exhibit 7.1.1 provides an overview of the Survey File segments and their inclusion of Community-only respondents, Facility-only respondents, or both types of respondents.

8.2.1 Sort Order for Merging the Survey File LDS Segments

Sort order is often important to understand when data users are merging segments within or across LDS releases. Most LDS segments are sorted by BASEID. However, some are sorted on other fields to create appropriate and unique sort keys for matching and merging the data. See Exhibit 8.2.1 below.

Exhibit 8.2.1: Sort Order by Segment in the Survey File LDS

Segment	Sorted by
ASSIST	BASEID HLPRNUM
FACCHAR	BASEID RECADMN
HITLINE	BASEID PLANTYPE PLANNUM
INTERV	BASEID SEQNUM
MDS3	BASEID TRGT_DT A2300
OASIS	BASEID HHASMTID

The MCBS Research Claims are a subset of items from the claims available on the Chronic Conditions Warehouse (CCW). All research claims are sorted by BASEID and CLAIMID. The MCBS Claims Variable Crosswalk spreadsheet crosswalks the MCBS claims item (variable) names with the CCW item (variable) names.

Item (variable) names are listed in alphabetical order. MCBS Research Claims have a unique and de-identified BASEID and CLAIMID, so that these cannot be linked back to the original claims.

The full descriptions of the items on the MCBS Research Claim can be found on the public facing CCW Claims Data Dictionary, located at: <https://www2.ccwdata.org/documents/10280/19022436/codebook-ffs-claims.pdf>

8.3 Claims Files

The fixed-length claims (also known as the research claims or FFS claims) are abbreviated versions of the full claim record layout. Each claim type has a subset of variables selected for their relevancy to data analysis of that service. Additionally, institutional claim types have a corresponding revenue center file that links back to the claim-level data file through a unique claim identifier. The Research Claims are provided as SAS files and as CSV files.

MCBS data can be linked to Medicare Part A and Part B claims data for beneficiaries who participated in the MCBS. MCBS data cannot be linked to electronic medical records, or to any other records that record lab values or physiologic data.

Starting with 2021, the Survey File LDS will include five years of research claims. This means that, depending on their original enrollment date and enrollment type (Medicare FFS vs. Medicare Advantage), a beneficiary could have up to five years of claims included with their MCBS data. This update is reflected in a change to the claims file names (e.g., the 2022 DME claims file is named "DME_18_22" instead of just "DME").

8.3.1 Utilization Detail Records

Core Content

The following rules were used to select claims records for the Claims files.

1. Inpatient claims were included if the discharge or "through" date fell on or after January 1, 2022 and on or before December 31, 2022.
2. Skilled nursing facility claims were included if the admission or "from" date fell on or after January 1, 2022 and on or before December 31, 2022.
3. Home health agency and outpatient facility claims were included if the "through" date fell on or after January 1, 2022 and on or before December 31, 2022.
4. Hospice claims were included if the admission or "from" date fell on or after January 1, 2022 and on or before December 31, 2022.
5. Physician or supplier claims were included if the latest "service thru" date fell on or after January 1, 2022 and on or before December 31, 2022.
6. Durable medical equipment (DME) claims were included if the latest "service thru" date fell on or after January 1, 2022 and on or before December 31, 2022.

A total of 5,531 (about 40 percent) of the 2022 survey respondents did not use Medicare reimbursed services in an FFS setting in 2022; consequently, there are no claims records for them in this file. These individuals may have used no services at all, services only in a managed care plan, or services provided by a payer other than Medicare.²³ For the other 8,448 individuals in the sample, the MCBS has captured claims meeting the date criteria, processed and made available by CMS through June 2023.²⁴

²³ The HITLINE segment provides data on types of insurances, the coverage eligibility timeline, and the source information for the coverage use of services (i.e., Medicare Administrative enrollment data and/or survey data). The ACCESSCR and ACCSSMED segments also provide self-reported data on access and satisfaction with visits. See the Data File Notes section of this document for more information on the contents of these segments.

²⁴ Note that claims "mature" through the midpoint of the following calendar year. That is, 2022 claims were pulled from CMS' administrative data after June 2023 to ensure that the 2022 claims had been finalized.

9. DATA FILE DOCUMENTATION

9.1 LDS Contents

In addition to the data, CMS provides technical documentation with the following resources for data users:

- Codebooks
- Questionnaires
- Data files (SAS, CSV)
- Research claims (SAS, CSV)
- Format control files
- Sample SAS code to apply the formats and labels for researchers not using SAS

9.2 LDS Components

9.2.1 Codebooks

Codebooks are included with each data release and serve as the key resource for comprehensive information on all variables within a data file. The codebooks list the variables in each of the segments, the possible values, and unweighted frequencies. For variables that are associated with items in the MCBS Questionnaire, the item number and item text are provided.

The information provided within each Codebook is as follows:

Variable: The Codebook contains the variable names associated with the final version of the data files. Certain conventions apply to the variable names. All variables that are preceded by the character "D_", such as D_ERVIST, are derived variables. Variables preceded by the character "H_", such as H_DOB, come from CMS administrative source files.

Format Name: This column identifies the format name associated with the variable in the SAS dataset.

Frequency: This column shows unweighted frequency counts of values or recodes for each variable.

Question #: This column contains a reference to the questionnaire for direct variables, or to the source of derived variables. For example, the entry that accompanies the variable D_ERVIST in the Access to Care, Medical Appointments segment is "AC1." The first question in the Access to Care portion of the Community Questionnaire is the one referenced. This column will be blank for variables that do not relate to the questionnaire or to the CMS administrative source files, which are usually variables created to manage the data and the file.

Description/Label (variable label and codes): The variable label provides an explanation of the variable, which describes it more explicitly than would be possible in only eight letters. For coded variables, all of the possible values of the variable appear in lines beneath that explanation. Associated with each possible value (in the column labeled "Frequency") is a count of the number of times that the variable had that value, and, under the column labeled "Label", a short format expanding on the coded value.

BASEID: The BASEID is the unique identifier assigned to each beneficiary. This identifier can be used to link data across the survey files.

Survey Year: The Survey Year of interest is included as a variable on the file.

Version Number: Files may be re-released due to needed updates, which will be noted by the version number variable.

Note: Each variable may be followed by a statement that describes when a question was not asked, resulting in a missing variable. Questions were not asked when the response to a prior question or other information gathered earlier in the interview would make them inappropriate. For example, respondents who indicated that they never smoked (Community interview, question HFG1) would not be asked if they currently smoke (question HFG2). Notes also describe important information about the variable. For variables added to the survey recently, the first year of administration is also listed in the note.

Many questions were written to elicit simple “Yes” or “No” answers, or to limit responses to one choice from a list of categories. In other questions, the respondent was given a list of responses and instructed to select all responses that applied. In these cases when the question was a “select all that apply” item, each of the responses is coded “Indicated”/“Selected” or “Not Indicated”/“Not Selected.”

9.2.2 Questionnaires

Data users can view the Questionnaires for each data year along with the questionnaire variable names and question text on the MCBS website at: <https://www.cms.gov/data-research/research/medicare-current-beneficiary-survey/questionnaires>.

9.2.3 Data User Resources

CMS provides technical assistance to researchers interested in using MCBS data and provides free consultation to users interested in obtaining these data products and using these data in research. Users can email MCBS@cms.hhs.gov with questions regarding obtaining or using the data.

9.3 Data Edits and Imputation

9.3.1 Data Edits

A series of checks and edits are conducted to ensure the accuracy, completeness, and reasonableness of data within each data file. Any structural issues are addressed during either data extraction or data cleaning.

Logic checks verify that the questionnaire worked as expected, particularly with respect to questionnaire routing. Errors identified during logic checking are addressed with two categories of data edits: flagging values that were incorrectly skipped and setting incorrectly populated values to null to indicate a valid missing value.

Additional checks identify unreasonable values that are not explicitly disallowed by the questionnaire (e.g., male beneficiaries reporting female-only conditions, such as cervical cancer). After investigation, such values are then addressed with global edits. The MCBS also conducts consistency checks to identify scenarios where respondents report inconsistent information (e.g., indicating that one is Medicaid eligible due to a certain condition, but not reporting having that same condition when asked about health status). Based on a thorough data review, these types of errors are addressed with edits during data cleaning.

Certain conventions are used in coding all variables to distinguish between questions that beneficiaries would not or could not answer and questions that were not asked. These conventional codes are depicted in Exhibit 9.3.1.

Exhibit 9.3.1: Data Review and Editing Codes

Value	Format	Meaning
.	INAPPLICABLE	Valid missing, inapplicable, a valid skip, missing with no expectation that a value should be present. Missing is '.' in numeric variables and blank in character variables.
.R	REFUSED	Valid missing, refused survey response
.D	DON'T KNOW	Valid missing, don't know survey response
.N	INVALID SKIP	Invalid missing, not ascertained, an invalid skip, a response should be present but is not
.E	EDITING CODE	Editing code, extreme value, unreasonable or out of range survey response
.S	SUPPRESSED*	Valid value suppressed due to suppression guidelines applied to Area Deprivation Index (ADI) variables

*Code not applied to data collected by the Facility Instrument

9.3.2 Imputation

To compile the most accurate and complete LDS, there are several types of adjustments applied to the MCBS data that compensate for missing information. Although a variety of methods are used in making the adjustments, adjustments of all types are governed by some basic principles. Information reported by the survey respondent is retained, even if it is not complete, unless strong evidence suggests that it is not accurate. When information is not reported during the interview, Medicare claims data and administrative data are the first choice as a source of supplementary, or in some cases, surrogate information.

There are several techniques for handling cases with missing data. One option is to impute the missing data. This can be done in such a way as to improve univariate tabulations, but techniques that retain correlation structure for multivariate analyses are extremely complex. For more discussion of imputation, see Kalton and Kasprzyk.²⁵

The MCBS imputes income when income data are missing. Using the hot deck imputation method, the MCBS first imputes whether an income source exists (such as Social Security). If the income source exists, then the amount earned was imputed. A flag was created for each imputed variable indicating whether the corresponding value was imputed.

The 2022 Income and Assets imputation used IAQ data reported in 2023, as the 2023 IAQ asks about total income in the prior year (2022). The MCBS imputed different sets of variables for respondents to the 2023 IAQ and for the 2022 ever enrolled respondents who did not complete the 2023 IAQ. For the first group, the MCBS imputed a selection of variables from the 2023 IAQ. These included probe variables, which are indicators of whether the beneficiary and/or the spouse/partner had income or asset items, and amount variables, which give the amount of the income or asset items that the beneficiary and/or the spouse/partner had. For the second group, which includes beneficiaries living in a facility, only the amount of total income was imputed.

Beginning with 2022, the HITLINE segment includes imputed premiums for beneficiaries who did not provide complete premium information about their private plan and the MAPLANQX segment includes the imputed premiums for MA plans.

²⁵ Graham Kalton and Daniel Kasprzyk, "The Treatment of Missing Survey Data," *Survey Methodology* 12, no. 1 (1986): 1-16.

The MCBS created one imputation flag for each imputed variable. For the probes, only the hot deck imputation method was used, so the imputation flags indicate whether the probe was imputed or not. For the amounts, the MCBS used a variety of imputation methods. The imputation flags indicate whether the amount was not imputed, imputed by the hot deck method, imputed by the carry forward method, or imputed by data edits. The imputation used information from the Income and Assets and Facility Assessments Survey File segments and demographic information from the Beneficiary Demographics and Household Characteristics segments.

Using information from the Cost Supplement File segments and Medicare claims data, the MCBS imputed missing payer and payment information for medical events reported in 2022. For beneficiaries living in a facility, medical event data are provided only from Medicare claims data. The MCBS first imputed whether a payer, such as an insurance plan, paid for a particular event. If the payer paid, then the amount paid was imputed next. Imputation was performed using the hot deck imputation method, and a flag was created for each imputed variable indicating whether the corresponding cost value was imputed.

Beginning with 2019, the MA encounter data were utilized to improve estimation of medical events and costs for beneficiaries enrolled in MA. Beginning with 2020, these MA encounter data adjustments were improved to better reflect age and general health-related differences.

9.4 Weighting

9.4.1 Preparing Statistics (Using the Full Sample Weights)

Two types of weights are provided, cross-sectional weights and longitudinal weights. Cross-sectional weights apply to the entire file of all beneficiaries who completed an interview, either Community or Facility. Cross-sectional weights are available for the Survey File and the Cost Supplement File in each data year.

The data user may choose to conduct analyses of the Survey File data alone or use the Cost Supplement data to conduct joint analyses of both survey and cost and utilization data. Exhibit 9.4.1 provides an overview of the weights for the 2022 Survey File and Cost Supplement File. For analysis of Survey File data, there are two populations of inference that can be obtained using two distinct weights. The ever enrolled Survey File weight is greater than zero for all beneficiaries in the Survey File. This weight segment is EVRWGTS, and the name of the weight is EEYRSWGT. The sum of this weight represents the population of beneficiaries who were entitled and enrolled in Medicare for at least one day at any time during the calendar year.

The continuously enrolled Survey File weight is greater than zero for the subset of beneficiaries in the Survey File who were continuously enrolled in Medicare from January 1, 2022, through completion of their fall interview. This weight segment is CENWGTS, and the weight is named CEYRSWGT. The population represented by the sum of this weight is the continuously enrolled population of Medicare beneficiaries who were enrolled from the first of the year through the Fall 2022.²⁶ Users should use the continuously enrolled Survey File weight (CEYRSWGT) for time series analysis of survey data across years.

Analyses of the Cost Supplement File data should be done with the Cost Supplement weight, which represents an ever enrolled population of Medicare beneficiaries enrolled in Medicare on at least one day at any time in 2022. To define the population, the MCBS creates a calendar history of a beneficiary's MCBS interviews. A number of eligibility checks are run against this calendar history to identify beneficiaries who met eligibility requirements for inclusion in the survey data for the calendar year, either because they were interviewed for a full year or interviewed until death or loss of Medicare entitlement. Beneficiaries who pass these eligibility

²⁶ This is identical to the historical Access to Care (ATC) cross-sectional weight that was available in previous years, 1991-2013.

checks become the population eligible for the Cost Supplement ever enrolled weight and the prescription medicine data files.

The Cost Supplement weights segment is named CSEVRWGT. The population represented by the sum of this weight is identical to the population represented by the sum of the ever enrolled Survey File weight, but it is populated for a smaller subset of respondents with complete cost and utilization data. Users wishing to conduct joint analysis of both Survey File and Cost Supplement File data should use the Cost Supplement File weights.

The weights mentioned above for the calendar year 2022 are full-sample weights. The term “full-sample” distinguishes these weights from the replicate weights used for variance estimation, as discussed in the Section 9.6: Variance Estimation. Additional information on using the weights is available in the file-specific MCBS Data User's Guide documents that accompany each data file release.

Longitudinal weights allow for the study of respondents across data years. The following longitudinal weights are provided with the 2022 Survey File and Cost Supplement LDS's.²⁷

Survey File Two-Year Longitudinal Weights (LNG2WGTS): Two-year longitudinal weights apply to respondents who completed fall round interviews in the current and the preceding year. This set of weights can be used to study data trends over a two-year period and are populated only for members of the 2019, 2020, 2021 panels who had 2021 and 2022 Survey File data and were continuously enrolled for two years. The population represented by these weights is the population of beneficiaries enrolled on or before 1/1/2021 and surviving and entitled as of completion of the Fall 2022 interview. By applying these weights to data in the current and preceding year, users will be able to estimate change among the Medicare population who were alive for the full two-year period.

Survey File Three-Year Longitudinal Weights (LNG3WGTS): Three-year longitudinal weights apply to respondents who completed fall round interviews in the current and the two preceding years. This set of weights can be used to study data trends over a three-year period and are populated only for members of the 2019 and 2020 panels who were continuously enrolled during all of the years 2020-2022 and had Survey File data in 2020 and 2022. The resulting weights represent the population of Medicare beneficiaries who enrolled on or before 1/1/2020 and were still alive and entitled as of completion of the Fall 2022 interview. By applying these weights to data in the current and the three preceding years, users will be able to estimate change among the Medicare population who were alive for the full three-year period.

Survey File Four-Year Longitudinal Weights (LNG4WGTS): Four-year longitudinal weights apply to respondents who completed fall round interviews in the current and the three preceding years. This set of weights can be used to study data trends over a four-year period and are populated only for members of the 2019 Panel who were continuously enrolled during all of the years 2019-2022. The resulting weights represent the population of Medicare beneficiaries who enrolled on or before 1/1/2019 and were still alive and entitled as of completion of the Fall 2022 interview. By applying these weights to data in the current and the three preceding years, users will be able to estimate change among the Medicare population who were alive for the full four-year period.

²⁷ Beginning with the 2016 LDS, the Survey File longitudinal weight names reflect the number of years the beneficiary was enrolled in Medicare (i.e., LNG2WGTS weights are referred to as 'two-year' rather than 'one-year' as they represent the population continuously enrolled for two years). This change was made to align the names of the longitudinal weights in the Survey File LDS with the naming convention used for the Cost Supplement LDS.

Cost Supplement Two-Year Longitudinal Weights (CSL2WGTS): The two-year longitudinal weights are populated for members of the 2019, 2020, and 2021 Panels who were ever enrolled in Medicare at any time during both 2021 and 2022 and provided utilization and cost data for both years.

Cost Supplement Three-Year Longitudinal Weights (CSL3WGTS): The three-year longitudinal weights are populated for members of the 2019 and 2020 Panels who were ever enrolled in Medicare at any time during 2020, 2021, and 2022, and provided utilization and cost data for all three years.

The Survey File longitudinal weights are for analysis of Survey File data. Data users cannot use the Survey File longitudinal weights with Cost Supplement data. Users who want to analyze Survey File data along with utilization and cost data in the Cost Supplement should limit analysis to cases with a positive Cost Supplement weight.

9.4.2 Special Topical Segment Weights

There are ten 2022 Survey File LDS segments (FOODINS, INCASSET, MCREPLNQ, RXMED, ACCSSMED, USCARE, PNTACT, CHRNPAIN, TELEMED, COVIDTOP) that have special non-response adjusted weights. These segments are referred to as Topical segments because most of them were traditionally sourced from the Topical questionnaire sections (see Section 5.2.4 for more information). **To generate estimates using these Topical segment data, on their own or merged with another Survey File segment, always use the special full-sample and replicate weights included in the Topical segment.** Do not use the weights that appear in the separate weight segments (CENWGTS, EVRWGTS).

The questionnaire sections (or specific items within questionnaire sections) that are weighted separately are fielded in the winter and summer rounds following the data year, and/or are not administered to proxy respondents. Exhibit 9.4.1 crosswalks these questionnaire sections and their corresponding Topical segments.

Exhibit 9.4.1: Crosswalk of 2022 Questionnaire Sections and LDS Segments with Topical Weights

Questionnaire Section	Questionnaire Type	Data Collection Round	Topical LDS Segment
Access to Care (ACQ)	Core	Winter 2023	ACCSSMED
Chronic Pain (CPQ)*	Topical	Summer 2023	CHRNPAIN
COVID-19 (CVQ)	Topical	Winter 2023, Summer 2023	COVIDTOP
Income and Assets (IAQ)	Core	Summer 2023	INCASSET
Income and Assets (IAQ) – Food Insecurity items	Topical	Summer 2023	FOODINS
Knowledge and Decision Making (KNQ)	Topical	Winter 2023	MCREPLNQ
Satisfaction with Care (SCQ) – Patient Activation items*	Core	Fall 2022	PNTACT
Drug Coverage (RXQ)	Topical	Summer 2023	RXMED
Telemedicine (TLQ)	Core	Winter 2023	TELEMED
Usual Source of Care (USQ)	Core	Winter 2023	USCARE

*CPQ and the Patient Activation items in SCQ are only administered to non-proxy respondents.

There are three sets of full-sample and replicate weights for each Topical segment, one based on the 2022 Survey File ever enrolled population, one based on the 2022 Survey File continuously enrolled population, and one based on the 2022 Cost Supplement ever enrolled population. These weights may be used to conduct joint analyses of Topical segment data, Survey File data, and Cost Supplement File data. Note that counts of cases with positive Topical weights may vary within the data year and may change across years due to response rates, sample sizes, and fielding methods. The Topical weights account for these changes.

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

The Topical weights that are described as "Survey File ever enrolled" weights (e.g., KNSEWT, INSEWT) correspond to the Survey File ever enrolled population and can be used to conduct analyses of the Topical data as representing the ever enrolled population and in conjunction with other Survey File data. The Topical weights that are described as "Survey File continuously enrolled" weights (e.g., KNSCWT, INSCWT) correspond to the Survey File continuously enrolled population and can be used to conduct analyses of the Topical data as representing the continuously enrolled population and in conjunction with other Survey File data. The Topical weights that are described as "Cost Supplement ever enrolled" weights (e.g., KNCEWT, INCEWT) correspond to the Cost Supplement ever enrolled population and can be used to conduct analyses of the Topical data as representing the ever enrolled population and in conjunction with Cost Supplement data. Weights corresponding to the Cost Supplement File continuously enrolled population are not available for the Topical data. Because the Cost Supplement is available for a smaller subset of the Survey File population, for each Topical segment the number of beneficiaries with a continuously enrolled Survey File Topical weight is larger than the number of beneficiaries with an ever enrolled Cost Supplement File Topical weight.

The Topical segments, weights, and weight names are listed in Exhibit 9.4.2. Please also see the forthcoming *2022 MCBS Methodology Report* for additional information on the composition and derivation of the Topical weights.

Prefixes for the weights changed slightly in 2018 to accommodate the additional new population and make the population clearer to the data users.

Exhibit 9.4.2: 2022 MCBS Data Files Summary of Weights

Limited Data Set	Description	Segment	Full-Sample Weight	Replicate Weights	Population
Survey File	Continuously Enrolled Cross-Sectional Weights	CENWGTS	CEYRSWGT	CEYRS001-CEYRS100	Continuously enrolled from 1/1/2022 through the fall of 2022
Survey File	Ever Enrolled Cross-Sectional Weights	EVRWGTS	EEYRSWGT	EEYRS001-EEYRS100	Ever enrolled for at least one day at any time during 2022
Survey File	Continuously Enrolled Two-Year Longitudinal Weights	LNG2WGTS	L2YRSWGT	L2YRS001-L2YRS100	Continuously enrolled from 1/1/2021 through the fall of 2022
Survey File	Continuously Enrolled Three-Year Longitudinal Weights	LNG3WGTS	L3YRSWGT	L3YRS001-L3YRS100	Continuously enrolled from 1/1/2020 through the fall of 2022
Survey File	Continuously Enrolled Four-Year Longitudinal Weights	LNG4WGTS	L4YRSWGT	L4YRS001-L4YRS100	Continuously enrolled from 1/1/2019 through the fall of 2022
Cost Supplement File	Ever Enrolled Cross-Sectional Weights	CSEVRWGT	CSEVRWGT	CSEVR001-CSEVR100	Ever enrolled for at least one day at any time during 2022
Cost Supplement File	Two-Year Longitudinal Weights	CSL2WGTS	CSL2YWGT	CSL2Y001-CSL2Y100	Enrolled at any time during both 2021 and 2022
Cost Supplement File	Three-Year Longitudinal Weights	CSL3WGTS	CSL3YWGT	CSL3Y001-CSL3Y100	Enrolled at any time during each of 2020, 2021, and 2022
Survey File Topical Section	KNQ Survey File Ever Enrolled	MCREPLNQ	KNSEWT	KNSE1-KNSE100	Ever enrolled in 2022 and still alive, entitled, and not living in a facility in Winter 2023

Limited Data Set	Description	Segment	Full-Sample Weight	Replicate Weights	Population
Survey File Topical Section	KNQ Survey File Continuously Enrolled	MCREPLNQ	KNSCWT	KNSC1-KNSC100	Continuously enrolled in 2022 and still alive, entitled, and not living in a facility in Winter 2023
Survey File Topical Section	KNQ Cost Supplement Ever Enrolled	MCREPLNQ	KNCEWT	KNCE1-KNCE100	Ever enrolled in 2010 and still alive, entitled, and not living in a facility in Winter 2023
Survey File Topical Section	ACQ Survey File Ever Enrolled	ACCSSMED	ACSEWT	ACSE1-ACSE100	Ever enrolled in 2022 and still alive, entitled, and not living in a facility in Winter 2023
Survey File Topical Section	ACQ Survey File Continuously Enrolled	ACCSSMED	ACSCWT	ACSC1-ACSC100	Continuously enrolled in 2022 and still alive, entitled, and not living in a facility in Winter 2023
Survey File Topical Section	ACQ Cost Supplement Ever Enrolled	ACCSSMED	ACCEWT	ACCE1-ACCE100	Ever enrolled in 2022 and still alive, entitled, and not living in a facility in Winter 2023
Survey File Topical Section	USQ Survey File Ever Enrolled	USCARE	USSEWT	USSE1-USSE100	Ever enrolled in 2022 and still alive, entitled, and not living in a facility in Winter 2023
Survey File Topical Section	USQ Survey File Continuously Enrolled	USCARE	USSCWT	USSC1-USSC100	Continuously enrolled in 2022 and still alive, entitled, and not living in a facility in Winter 2023
Survey File Topical Section	USQ Cost Supplement Ever Enrolled	USCARE	USCEWT	USCE1-USCE100	Ever enrolled in 2022 and still alive, entitled, and not living in a facility in Winter 2023

Limited Data Set	Description	Segment	Full-Sample Weight	Replicate Weights	Population
Survey File Topical Section	TLQ Survey File Ever Enrolled	TELEMED	TMSEWT	TMSE1-TMSE100	Ever enrolled in 2022 and still alive, entitled, and not living in a facility in Winter 2023
Survey File Topical Section	TLQ Survey File Continuously Enrolled	TELEMED	TMSCWT	TMSC1-TMSC100	Continuously enrolled in 2022 and still alive, entitled, and not living in a facility in Winter 2023
Survey File Topical Section	TLQ Cost Supplement Ever Enrolled	TELEMED	TMCEWT	TMCE1-TMCE100	Ever enrolled in 2022 and still alive, entitled, and not living in a facility in Winter 2023
Survey File Topical Section	IAQ Survey File Ever Enrolled	INCASSET	INSEWT	INSE1-INSE100	Ever enrolled in 2022 and still alive, entitled, and not living in a facility in Summer 2023
Survey File Topical Section	IAQ Survey File Continuously Enrolled	INCASSET	INSCWT	INSC1-INSC100	Continuously enrolled in 2022 and still alive, entitled, and not living in a facility in Summer 2023
Survey File Topical Section	IAQ Cost Supplement Ever Enrolled	INCASSET	INCEWT	INCE1-INCE100	Ever enrolled in 2022 and still alive, entitled, and not living in a facility in Summer 2023
Survey File Topical Section	IAQ Survey File Ever Enrolled	FOODINS	FDSEWT	FDSE1-FDSE100	Ever enrolled in 2022 and still alive, entitled, and not living in a facility in Summer 2023
Survey File Topical Section	IAQ Survey File Continuously Enrolled	FOODINS	FDSCWT	FDSC1-FDSC100	Continuously enrolled in 2022 and still alive, entitled, and not living in a facility in Summer 2023

Limited Data Set	Description	Segment	Full-Sample Weight	Replicate Weights	Population
Survey File Topical Section	IAQ Cost Supplement Ever Enrolled	FOODINS	FDCEWT	FDCE1-FDCE100	Ever enrolled in 2022 and still alive, entitled, and not living in a facility in Summer 2023
Survey File Topical Section	PAQ Survey File Enrolled	PNTACT	PASEWT	PASE1-PASE100	Ever enrolled for at least one day at any time during 2022
Survey File Topical Section	PAQ Survey File Continuously Enrolled	PNTACT	PASCWT	PASC1-PASC100	Continuously enrolled from 1/1/2022 through the fall of 2023
Survey File Topical Section	PAQ Cost Supplement Ever Enrolled	PNTACT	PACEWT	PACE1-PACE100	Ever enrolled for at least one day at any time during 2022
Survey File Topical Section	RXQ Survey File Ever Enrolled	RXMED	RXSEWT	RXSE1-RXSE100	Ever enrolled in 2022 and still alive, entitled, and not living in a facility in Summer 2023
Survey File Topical Section	RXQ Survey File Continuously Enrolled	RXMED	RXSCWT	RXSC1-RXSC100	Continuously enrolled in 2022 and still alive, entitled, and not living in a facility in Summer 2023
Survey File Topical Section	RXQ Cost Supplement Ever Enrolled	RXMED	RXCEWT	RXCE1-RXCE100	Ever enrolled in 2022 and still alive, entitled, and not living in a facility in Summer 2023
Survey File Topical Section	CPQ Survey File Ever Enrolled	CHRNPAIN	CPSEWT	CPSE1-CPSE100	Ever enrolled in 2022 and still alive, entitled, and not living in a facility in Summer 2023
Survey File Topical Section	CPQ Survey File Continuously Enrolled	CHRNPAIN	CPSCWT	CPSC1-CPSC100	Continuously enrolled in 2022 and still alive, entitled, and not living in a facility in Summer 2023

Limited Data Set	Description	Segment	Full-Sample Weight	Replicate Weights	Population
Survey File Topical Section	CPQ Cost Supplement Ever Enrolled	CHRNPAIN	CPCEWT	CPCE1-CPCE100	Ever enrolled in 2022 and still alive, entitled, and not living in a facility in Summer 2023
Survey File Topical Section	CVQ Survey File Ever Enrolled	COVIDTOP	VSSEWT	VSSE1-VSSE100	Ever enrolled in 2022 and still alive, entitled, and not living in a facility in Summer 2023
Survey File Topical Section	CVQ Survey File Continuously Enrolled	COVIDTOP	VSSCWT	VSSC1-VSSC100	Continuously enrolled in 2022 and still alive, entitled, and not living in a facility in Summer 2023
Survey File Topical Section	CVQ Cost Supplement Ever Enrolled	COVIDTOP	VSCEWT	VSCE1-VSCE100	Ever enrolled in 2022 and still alive, entitled, and not living in a facility in Summer 2023

9.5 Using the Data

9.5.1 Merging Segments within 2022

Data users can merge segments within and/or across the Survey File and Cost Supplement File. Appendix C provides a hypothetical research question with sample SAS code for the construction of an analytic file using the 2022 Survey File LDS. For an example of how to merge data across the Survey File and Cost Supplement File LDS's, please see Appendix B.1: Using the Data of the *Data User's Guide: Cost Supplement File*. Note that although the MCBS data are nationally representative, they are not representative at the regional or state level and cannot be used to produce regional or state-level estimates. However, the data user can use the data to look for national trends across population groups.

9.6 Variance Estimation (Using the Replicate Weights)

9.6.1 Variables Available for Variance Estimation

In many statistical packages, the procedures for calculating sampling errors (e.g., variances, standard errors, margins of error) assume that the data were collected in a simple random sample. Procedures of this type are not appropriate for calculating the sampling errors of statistics based upon a stratified, unequal-probability, multi-stage sample such as the MCBS. Unless the complex nature of the MCBS is taken into account, estimates of the variance of a survey statistic may be biased downward.

The MCBS includes variables to obtain weighted estimates and estimated standard errors using either the Taylor-series linearization approach or balanced repeated replication (BRR) method, also known as Fay's method. There is both serial and intra-cluster correlation in the MCBS data, including: sampling second-stage units within primary sampling units; sampling beneficiaries with second-stage units; and repeated observations of the selected beneficiary across time. Researchers should use the BRR method of variance estimation to account for various correlations. For details on the strengths and weaknesses of the two variance estimation methods, please refer to Wolter.²⁸

To estimate variance using the balanced repeated replication method, a series of replicate weights are included in the 2022 Survey File release. As displayed in Exhibit 9.4.1 above, there are many types of full-sample weights, including those for cross-sectional analyses, longitudinal analyses, and analyses of Topical data. Each of these full-sample weights has a corresponding set of replicate weights. The replicate weights can be used to calculate standard errors of the sample-based estimates as described below. For the Survey File, the replicate cross-sectional weights are labeled CEYRS001 through CEYRS100 corresponding to the continuously enrolled weight CEYRSWGT, and EEYRS001 through EEYRS100 corresponding to the ever enrolled weight EEYRSWGT. These weights may be found on CENWGTS and EVRWGTS respectively. The Survey File replicate longitudinal weights are found on segments LNG2WGTS, LNG3WGTS, and LNG4WGTS.

The variables SUDSTRAT (sampling strata) and SUDUNIT (primary sampling unit) are used for variance estimation using the Taylor series linearization method. As with the BRR method, the Taylor series method can be applied for variance estimation using SUDSTRAT and SUDUNIT for any of the Core and Topical segments with full-sample weights. The Taylor series method does not employ the replicate weights that are included on the segments. For examples and guidance on using the Taylor series linearization method of variance estimation or the BRR method, please see Appendix B.

²⁸ Kirk Wolter, *Introduction to Variance Estimation* (Springer Science & Business Media, 2007).

9.6.2 Variance Estimation for Analyses of Single Year of MCBS

Most commercial software packages today include techniques to accommodate the complex design, either through Taylor-expansion type approaches or replicate weight approaches. Among these are R®, STATA®, SUDAAN®, and the complex survey procedures in SAS.

9.6.3 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.

9.7 Combining Multiple Years of Data

The MCBS is based on a rotating panel design, which allows for longitudinal analysis of up to four years when appropriate longitudinal weights are used. Multiple years of MCBS data can also be pooled to perform serial cross-sectional or pooled analysis. The appropriate method to combine data across years will depend on the analytic design of the study. Sample code is presented in Appendix B to demonstrate the steps involved in combining multiple years of data to perform two types of analysis: (1) Longitudinal analysis; (2) Pooled, cross-sectional analysis.

9.7.1 Longitudinal Analysis

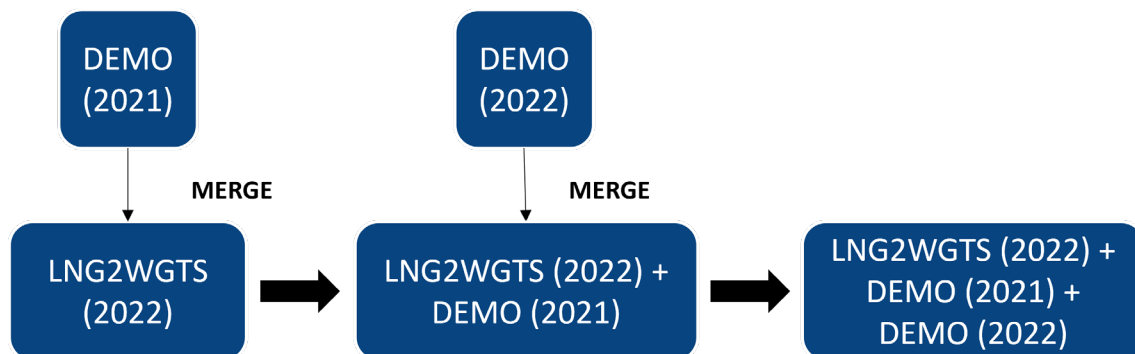
The study objective in longitudinal analysis is to assess changes over time for each sample person. The Survey File cross-sectional and longitudinal population definitions are consistent from year to year, so the data are comparable between years. The Cost Supplement cross-sectional population definition is also consistent and comparable from year to year.

Most longitudinal analyses require the data to be in long-format (i.e., repeated observations – each representing a calendar year the sample person was surveyed – are stored in a separate row for each sample person). To construct a longitudinal analytic dataset, the first step is to use the appropriate longitudinal weights file. For example, as shown in Exhibit 9.7.1, to assess changes over time beneficiaries who have been in the sample for at least two years – from CY2021 to CY2022 – the two-year longitudinal weights (i.e., one-year “backward longitudinal weights”) (LNG2WGTS) should be used.

Variables from current year files representing the outcome of interest should then be merged with the current year's longitudinal weights file. While merging, all observations in the weights file should be preserved. Next,

the same variables from the prior year's files should be merged with the current year's longitudinal weights file.

Exhibit 9.7.1: Constructing a Longitudinal Analytic File



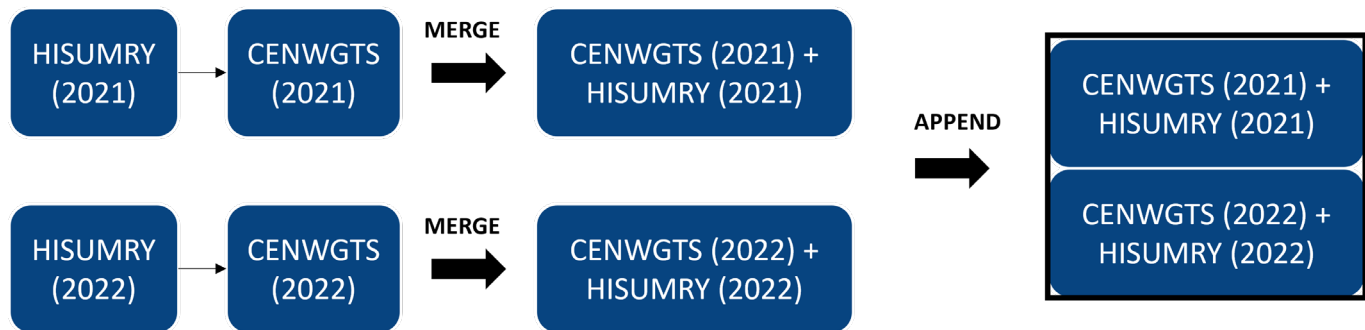
Variance estimation for longitudinal analysis (using replicate weights)

Just as there are full-sample longitudinal weights, there are corresponding sets of replicate weights. The replicate weights included in the longitudinal weights data files can be used to calculate standard errors of the sample-based estimates. The first set of replicate longitudinal weights is labeled L2YRS001 through L2YRS100 and may be found on the two-year longitudinal weights file (LNG2WGTS). The second set of replicate longitudinal weights is labeled L3YRS001 through L3YRS100 and may be found on the three-year longitudinal weights file (LNG3WGTS). The third set of replicate longitudinal weights in the Survey File LDS is labeled L4YRS001 through L4YRS100 and may be found on the four-year longitudinal weights file (LNG4WGTS).

For additional guidance, see the *MCBS Advanced Tutorial on Longitudinal Analysis*: <https://www.cms.gov/data-research/research/medicare-current-beneficiary-survey/data-briefs-tutorials>.

9.7.2 Repeated Cross-Sectional or Pooled Analysis

Multiple years of MCBS data can be pooled to perform serial cross-sectional or pooled analysis. Repeated cross-sectional analysis is used for analyzing changes in the Medicare population as a whole over time, while a pooled analysis will produce single cross-sectional estimates that cover a period of multiple years, usually to increase total sample sizes. In contrast, the longitudinal analysis described earlier is used to analyze beneficiary-level changes over time. Pooled data analysis yield estimates that are in effect a moving average of nationally representative year-specific estimates. The pooled estimates can be interpreted as being representative of the midpoint of the calendar year of the pooled period. Exhibit 9.7.2 demonstrates the steps involved in constructing a repeated cross-sectional or pooled analytic dataset using CY2021 and CY2022 data. For each year in the study, variables representing the outcome of interest should then be merged with the cross-sectional weights file. While merging, all observations in the weights file should be preserved. Next, the year-specific files are appended to produce the analytic dataset.

Exhibit 9.7.2: Constructing a Repeated Cross-Section or Pooled Analytic File**Variance estimation for repeated cross-sectional or pooled analysis (using replicate weights)**

Due to the rotating-panel and multistage-sampling design of the MCBS, there is both serial and intra-cluster correlation in the data when pooling multiple years of data. When conducting a pooled analysis, using the balanced half-sample method (also known as the balanced repeated replication, or BRR, method) of variance estimation throughout appropriately accounts for the various correlations due to sampling second-stage units within primary sampling units, sampling beneficiaries within second-stage units, and repeated observations of the selected beneficiary across time. The replicate cross-sectional weights are labeled CEYRS001 through CEYRS100 and can be found in each year's cross-sectional weights file (CENWGTS). A pooled analysis of this type can also be applied to longitudinal weights (LNG2WGTS, LNG3WGTS, LNG4WGTS).

When conducting a repeated cross-sectional analysis to compare between two years, the difference or net change in a population characteristic is often of interest. In this type of analysis, a point estimate of year-to-year difference is straightforward to calculate; simply take difference between the two individual annual cross-sectional estimates. Each cross-sectional estimate included in the comparison can be calculated using the full-sample weights included in that year's data release.

Calculating variance and standard error estimates of net change is more complicated because of correlation between the two annual data sets. Correlation is present because many beneficiaries are retained from one year to the next, and because the same set of PSUs and SSUs are used for each year. We refer to these types of correlation as serial and intra-cluster correlation, respectively.

To estimate the variance of net change estimates, the researcher may rely on a program such as SAS or calculate them directly in their own custom program using a closed formula.

For additional guidance, see the *MCBS Advanced Tutorial on Pooled Cross-Sectional Analysis*: <https://www.cms.gov/data-research/research/medicare-current-beneficiary-survey/data-briefs-tutorials>.

SAS Method

In SAS, point estimates of year-to-year differences, in addition to corresponding estimates of standard errors, can be generated using PROC SURVEYREG. To use this method, first concatenate the two annual datasets by stacking them together vertically, each including its corresponding set of weights, and define a YEAR variable to indicate which data year each of the two files represents. From this concatenated dataset, the example SAS code below will output estimates of the difference in estimates between the two years, using Cost Supplement weights as an example. The standard errors associated with these estimates are the desired estimated standard errors of the year-to-year net change.

```

PROC SURVEYREG VARMETHOD=BRR(FAY=.30);
  CLASS YEAR;
  MODEL variable = YEAR;
  LSMEANS YEAR / DIFF;
  WEIGHT CS1YRWGT;
  REPWEIGHT CS1YR001-CS1YR00100;
RUN;

```

This process can be repeated for any combination of variable and a complementary set of cross-sectional full-sample and replicate weights (e.g., Survey File continuously enrolled, Survey File ever enrolled, Cost Supplement ever enrolled weights).

Direct Method

The variance of a difference can also be calculated directly using the formula below, which a researcher can incorporate into a custom program for producing a variety of estimates of net change. This process does not require concatenating two annual files together, although programmatically it may be useful to do so. Let X_0^t be the cross-sectional estimate of the mean of population characteristic Y from year t using the full-sample weights from that year, and let $X_1^t, X_2^t, \dots, X_{100}^t$ be cross-sectional estimates of the same population mean from year t using each of the 100 corresponding replicate weights. Similarly, let $X_0^{t-1}, X_1^{t-1}, X_2^{t-1}, \dots, X_{100}^{t-1}$ be analogous estimates of the same population characteristic Y from year t-1, using the weights from year t-1. Next, define a set of difference variables as $D_0 = X_0^t - X_0^{t-1}$, $D_1 = X_1^t - X_1^{t-1}$, etc.

Then,

$$Var(D_0) = \frac{2.04}{100} \sum_{i=1}^{100} (D_i - D_0)^2$$

is an estimate of the variance of the estimate of net change from year t-1 to year t. The square root of this estimate is the estimated standard error.

10. DATA FILE NOTES

This section is a collection of information about various data fields present in the Survey File segments. The MCBS does not attempt to present information on every survey data field; rather, it concentrates its efforts on data fields where additional clarity or detail may be useful. The MCBS starts with information that is applicable globally, followed by specific information on individual segments, presented in the same sequence as the segments appear in the Codebook.

10.1 Global Information

10.1.1 BASEID

The BASEID key identifies the person interviewed. It is an 8-digit element, consisting of a unique, randomly assigned 7-digit number concatenated with a single-digit check digit.

LDS segments may vary in the number of BASEIDs. This variation may occur for several reasons. First, some segments include data from Community components and others from Facility components with different numbers of beneficiaries providing responses. Second, there are also differences in the number of beneficiaries by the specific round completed. Third, the use of ever enrolled or continuously enrolled weights in constructing the segments may result in differences.

10.1.2 Missing Values

Various special values indicate the reason why some data are missing, such as .R for "refused," .D for "don't know." See Exhibit 9.3.1 above for additional values.

10.1.3 Derived and Administrative Variables

Variables that were derived or created by combining two or more survey variables are preceded with the characters "D_", such as D_ERVIST. CMS may create or modify variables in order to recode data items (e.g., to protect the confidentiality of survey responses) or to globally edit some variables. Variables preceded by the character "H_", such as H_DOB, come from CMS administrative source files.

Several segments include variables indicating the length of time the beneficiary spent doing something, such as waiting in the hospital emergency room or waiting for an appointment. In the questionnaire, the length of time is stored in a continuous variable while the corresponding unit is stored in a categorical variable (e.g., hours, minutes, or hours and minutes; days, weeks, or months). These two variables are used to derive a single variable indicating the length of time in the most appropriate unit of time. For example, on the ACCSSMED segment, D_ERTIME contains the length of time spent waiting in the hospital emergency room in minutes while D_MDAPPT contains the length of time spent waiting for a doctor's appointment in days.

10.1.4 Initial Interview Variables

Some questions are asked in only two scenarios: 1) it is the case's Baseline (initial) interview or 2) it is the first time the case has crossed to a new component (e.g., the case crosses from the Community component to the Facility component for the first time). These "initial interview variables" are not asked again during subsequent interviews because the responses are not likely to change. Such questions include "Have you ever served in the armed forces?" and "What is the highest grade of school you ever completed?" To maximize the usefulness of this release as a cross-sectional file, these data are pulled forward from the Baseline interview or the first time the case was interviewed in a given component, as applicable. Variables that have been processed this way are listed in Appendix C.

10.1.5 Ever Variables

Many items in the MCBS ask respondents whether they have ever had certain experiences, such as ever being told they have a chronic condition, receiving a treatment, or doing a specific activity (such as ever accessing the official Medicare website). Such questions include "Have you ever been diagnosed with diabetes?" and "Have you smoked at least 100 cigarettes in your entire life?" Their responses are coded affirmatively if the respondent reports "yes" to having had that condition or experience.

These items are administered to respondents in certain scenarios. For select "ever" variables administered in the HFQ, there are different versions of each question, depending on whether a respondent is in the Incoming Panel sample or Continuing sample. These versions are combined into recoded variables to provide a complete picture of the response. All Incoming Panel sample respondents are asked if they have ever had certain conditions or experiences. Once a condition or experience is reported, the CAPI questionnaire logic retains that information for subsequent interviews. For variables that ask about conditions that cannot change after diagnosis, such as Alzheimer's, once an affirmative response is given, respondents are not asked again. However, if a negative response is given, respondents are asked annually thereafter if they had that specific condition or experience in the past year. For conditions that can change after diagnosis or can be reoccurring, such as high blood pressure, respondents are asked annually thereafter if they had that specific condition or experience in the past year. All data from a beneficiary from the current survey year and all previous years are used to determine whether the beneficiary has ever had a condition or experience.

"Ever" variables in the NAQ and KNQ are collected and processed in a similar manner to the HFQ "ever" variables, except that the NAQ and KNQ "ever" variables use only one version of each question (rather than two separate versions depending on beneficiary sample type).

For more information about "ever" variables pertaining to chronic conditions, see the data notes for the Chronic Conditions segment (CHRNCOND) in Section 10.3.5. For more information about "ever" variables pertaining to beneficiary knowledge about the Medicare open enrollment period and Medicare-covered expenses, see the data notes for the Medicare Plan Beneficiary Knowledge segment (MCREPLNQ) in Section 10.3.21. For more information about "ever" variables pertaining to beneficiary's nicotine and alcohol use, see the data notes for the Nicotine and Alcohol segment (NICOALCO) in Section 10.3.26.

10.1.6 Data Editing

Data are edited for consistency and to provide users with files that are easily used for analysis. Please see section 9.3.1 for additional detail on data edits.

10.1.7 Other Specify Questions

A subset of MCBS questionnaire items include closed ended responses with "other specify" options. These options allow respondents to provide answers that are not included in the existing code frame and are useful for questions with a wide range of potential responses (e.g., types of problems experienced during attempts to obtain care). In the event that an "other specify" option is selected, interviewers record actual responses verbatim. Verbatim responses are not released.

The MCBS programmatically identifies "other specify" responses that are sufficiently similar to existing code frame options and back codes responses into existing response option categories as appropriate. Often there will be more than one answer to a single question. In these cases, responses are recoded into several variables, all of which contain categorized data. Code lists are updated when necessary to incorporate responses that are frequently provided in "other specify" response options.

10.1.8 Interview Mode Indicator

MCBS data collection is multimode, with both in-person and phone interviewing. The Interview Characteristics (INTERV) segment includes a flag to indicate whether the interview was conducted in-person (INTMODE = 1), by telephone (INTMODE = 2), or using a combination of the two modes ("hybrid", INTMODE = 3). Data users should note potential mode effects when conducting analyses. Mode effects are discussed in further detail in the *2022 MCBS Methodology Report*.

10.1.9 Consistency with Medicare Program Statistics

In general, MCBS estimates may differ from Medicare program statistics using 100 percent administrative enrollment data. There are several reasons for the differences. The most important reason for the difference is that the administrative enrollment data may include people who are no longer alive. This may occur where people have entitlement, such as for Part A only, and receive no Social Security check. When field interviewers try to locate these beneficiaries for interviews, they establish the fact of these deaths. Unrecorded deaths may still be present on the Medicare Administrative enrollment data. The MCBS makes every effort to reconcile the survey information against the administrative data when possible. Other reasons, such as sampling error, may also contribute to differences between MCBS estimates and Medicare program statistics. Lastly, estimates may differ because Medicare program statistics adjust for partial enrollment. Medicare program estimates use a "person year" calculation where partial enrollment is counted as a fraction for the year. In contrast, the MCBS gives each beneficiary the same weight regardless of full or partial enrollment during the year, thus leading to differences in estimates using Medicare published statistics and MCBS data.

10.1.10 Do Administrative Data Override Survey-Reported Data?

In linking survey-reported and administrative data, the MCBS keep records from both sources to provide more complete data. Indicators in the file will usually specify if the information is survey-reported only, administrative data only, or both. Data that are only administrative are indicated as such in the data documentation and codebook.

10.2 Survey File Segment Information

Below is the information regarding each segment within the Survey File release, presented in alphabetical order. The notes have been organized into three main categories of information.

1. Core Content – a description of the main subject of the data.
2. Variable Definitions – definitions of derived variables and/or variables that require additional explanation regarding their construction. Note: The variables listed are not a comprehensive list of all variables in each segment. The Codebook provides information on all variables in each segment.
3. Special Notes – additional background information that data users may find helpful for constructing analyses.

10.3 Survey File Segment Descriptions

10.3.1 Access to Care (ACCESSCR)

10.3.1.1 Core Content

The Access to Care segment contains information from the HFQ section in the fall round. General questions are asked about the beneficiary's ability to access medical services. This segment also contains information on medical debt and the reasons beneficiaries cannot access the care they need.

10.3.1.2 Variable Definitions

Please see the Codebook for information regarding variables in this segment.

10.3.1.3 Special Notes

Respondents are asked why the beneficiary had trouble getting health care or scheduling a health care appointment in an open-ended format (e.g., "What were the reasons the doctor's office offered as an explanation for not scheduling an appointment with you?"). The respondents answer these questions in their own words, and interviewers select the response option(s) from a predefined code list that best matched the respondents' answer(s). These questions are select-all-that-apply so that respondents may provide multiple answers to each question, and each answer is stored in its own analytic variable.

If the respondent reports a reason that is not included in the predefined code list, the interviewer documents their response verbatim in an "other specify" variable that is not released. The "other specify" response is back coded as necessary into the predefined code list.

If the respondent reports that the beneficiary could not schedule an appointment because the doctor is not accepting new Medicare patients or the doctor does not accept Medicare at all, the respondent is then asked at variable OFFEXPLN whether the doctor's office explained why this is the case. If the doctor's office provided an explanation to the respondent, this explanation is recorded verbatim at OFFEXVB1 but not released.

10.3.2 Access to Care, Medical Appointments (ACCSSMED)

10.3.2.1 Core Content

The Access to Care, Medical Appointments segment contains information from the ACQ section and the emergency room, outpatient, medical provider, dental, vision, and hearing, and prescription medicine utilization sections asked in the winter round following the year of interest. General questions are asked about the beneficiary's access to all types of medical services and prescription medicines, the reasons for their visits, and the reasons for any forgone care or prescription medicines.

10.3.2.2 Variable Definitions

Please see the Codebook for information regarding variables in this segment.

10.3.2.3 Special Notes

Respondents are asked why the beneficiary did or did not receive different types of medical services or prescription medicines in an open-ended format (e.g., "What was the reason you saw the doctor?"). The respondents answer these questions in their own words, and interviewers select the response option(s) from a predefined code list that best matched the respondents' answer(s). These questions are select-all-that-apply so that respondents may provide multiple answers to each question, and each answer is stored in its own analytic variable.

If the respondent reports a reason that is not included in the predefined code list, the interviewer documents their response verbatim in an "other specify" variable that is not released. The "other specify" response is back coded as necessary into the predefined code list.

Special non-response adjustment weights are included in the file to account for survey non-response from the fall to the winter data collection period. Note that counts of cases with positive Topical weights may vary within the data year and may change across years due to response rates, sample sizes, and fielding methods.

The Topical weights account for these changes. Please see section 9.4.2 for information on using Topical weights.

10.3.3 Administrative Utilization Summary (ADMNUTLS)

10.3.3.1 Core Content

The Administrative Utilization Summary segment contains information on Medicare program expenditures and utilization taken directly from the Medicare Administrative enrollment data.

10.3.3.2 Variable Definitions

Except as noted otherwise, the variables in this segment are derived from summarizing data from CMS' Medicare Administrative enrollment data and the Medicare Administrative utilization and payment records. Administrative data available as of December 31, 2022 were summarized to create these data items.

H_HHASW: One or more home health agency (HHA) visits in the calendar year is indicated by a 1. This is triggered by a value (greater than zero) found in the home health visits field (H_HHVIS). Otherwise, the value for H_HHASW is 2.

H_HOSSW: One or more hospice bills in the calendar year is indicated by a 1. This is triggered by a value (greater than zero) found in the hospice Medicare payments (H_HOSPMT) field or the hospice stays (H_HOSSTY) field. Otherwise, the value for H_HOSSW is 2.

H_INPSW: One or more inpatient discharges in the calendar year is indicated by a 1. This is triggered by a value (greater than zero) found in the acute inpatient stays (H_ACTSTY) field or the other inpatient stays (H_OIPSTY) field. Otherwise, the value for H_INPSW is 2.

H_OUTSW: One or more outpatient visits in the calendar year is indicated by a 1. This is triggered by a value (greater than zero) found in the hospital outpatient visits (H_HOPVIS) field or hospital outpatient emergency room visits (H_HOP_ER) field. Otherwise, the value for H_OUTSW is 2.

H_PBSW: One or more Part B claims in the calendar year is indicated by a 1. This is triggered by a value (greater than zero) found in any of the following fields: H_PHYPMT, H_PHYEVT, H_PB_DEV, H_PB_DRG, H_PB_OTH, H_PB_OEV, H_DME EVT, H_DMEPMT, H_TSTEVT, H_TSTPMT, H_ANEVT, H_ANEPMT, H_ASCEVT, H_ASCPMT, H_DIAEVT, H_DIAPMT, H_EMEVT, H_EMPMT, H_IMGEVT, H_IMGPMPT, and H_PTBRMB. Otherwise, the value for H_PBSW is 2.

H_SNFSW: One or more skilled nursing facility (SNF) admissions in the calendar year is indicated by a 1. This is triggered by a value (greater than zero) found in any of the following fields: H_SNFPMT, H_SNFSTY, H_SNFDAY. Otherwise, the value for H_SNFSW is 2.

H_PTARMB: Total Part A reimbursement in the calendar year. It is a sum of calendar year reimbursements for HHA Part A, Hospice, Inpatient, and SNF. The CLM_PMT_AMT field is selected for each claim type in preparing this calculation. The CLM_VAL_CD = "64" is used to determine HHA Part A.

H_PTBRMB: Total Part B reimbursement in the calendar year. It is a sum of calendar year reimbursements for HHA Part B, Physician, and Outpatient. The CLM_PMT_AMT field is selected for each claim type in preparing this calculation. The CLM_VAL_CD = "65" is used to determine HHA Part B. "Physician" as noted in the "sum" statement above consisted of BCARRIER_CLAIMS and DME_CLAIMS.

H_ACTPMT: Acute Inpatient Medicare Payments is the sum of the Medicare claim payment amounts (CLM_PMT_AMT from each source claim) in the acute inpatient hospital setting in the calendar year. To obtain the total acute hospital Medicare payments, take this variable and add in the annual per diem payment amount (H_ACTMPT + H_ACTPRD).

H_ACTPRD: Acute Inpatient Hospital Pass-thru Per Diem Payments is the sum of all the pass-through per diem payment amounts (CLM_PASS_THRU_PER_DIEM_AMT from each source claim) in the acute inpatient hospital setting for the calendar year. Medicare payments are designed to include certain "pass-through" expenses such as capital-related costs, direct medical education costs, kidney acquisition costs for hospitals that are renal transplant centers, and bad debts. This variable is the sum of all the daily payments for pass-through expenses. It is not included in the Medicare Payment amount (H_ACTPMT). To determine the total Medicare payments for acute hospitalizations for the beneficiary, this field must be added to the total Medicare payment amount for acute inpatient hospitalizations (H_ACTPMT+ H_ACTPRD).

H_ACTSTY: Acute Inpatient Stays is the count of acute inpatient hospital stays (unique admissions, which may span more than one facility) for the calendar year. An acute inpatient stay is defined as a set of one or more consecutive acute inpatient hospital claims where the beneficiary is only discharged on the most recent claim in the set. If a beneficiary is transferred to a different provider, the acute stay is continued even if there is a discharge date on the claim from which the beneficiary was transferred.

H_ACTDAY: Acute Inpatient Medicare Covered Days is the count of Medicare covered days in the acute inpatient hospital setting for the calendar year.

H_ACTBPT: Acute Inpatient Beneficiary Payments is the sum of Medicare coinsurance and deductible payments in the acute inpatient hospital setting for the calendar year. The total acute hospitalization beneficiary payments are calculated as the sum of the beneficiary deductible amount and coinsurance amount for all acute inpatient claims where the CLM_PMT_AMT ≥ 0 .

H_IP_ER: Inpatient Emergency Room Visits is the count of emergency department (ED) claims in the inpatient setting for the year. The revenue center codes indicating emergency room use were 0450, 0451, 0452, 0456, and 0459.

H_OIPPMT: Other Inpatient Hospital Medicare Payments is the sum of the Medicare claim payment amounts (CLM_PMT_AMT from each source claim) in the other inpatient (OIP) settings for the calendar year. To obtain the total OIP Medicare payments, take this variable and add in the annual per diem payment amount (H_OIPPMT + H_OIPPRD). These OIP claims are a subset of the IP claims consisting of data from IP settings such as long-term care hospitals, inpatient psychiatric facilities, inpatient rehabilitation facilities, and other types of IP facilities such as children's hospitals or cancer centers.

H_OIPPRD: Other Inpatient Pass-thru Per Diem Payments is the sum of all the pass-through per diem payment amounts (CLM_PASS_THRU_PER_DIEM_AMT from each source claim) in the OIP setting for the calendar year. This variable is the sum of all the daily payments for pass-through expenses. It is not included in the Medicare payment amount (H_OIPPMT). To determine the total Medicare payments for other non-acute hospitalizations for the beneficiary, this field must be added to the total Medicare payment amount for other hospitalizations (H_OIPPMT + H_OIPPRD).

H_OIPSTY: Other Inpatient Stays is the count of hospital stays (unique admissions, which may span more than one facility) in the non-acute inpatient setting for the calendar year. A non-acute inpatient stay is defined as a set of one or more consecutive non-acute inpatient claims where the beneficiary is only discharged on the most recent claim in the set.

H_OIPDAY: Other Inpatient Hospital Covered Days is the count of covered days in the non-acute inpatient hospital setting for the calendar year. This variable equals the sum of the CLM_UTLZTN_DAY_CNT variables on the source claims.

H_OIPBPT: Other Inpatient Beneficiary Payments is the sum of Medicare coinsurance and deductible payments in the nonacute inpatient hospital setting for the year. The total OIP beneficiary payments are calculated as the sum of NCH_BENE_IP_DDCTBL_AMT and NCH_BENE_PTA_COINSRNC_LBLTY_AM for all relevant claims where the CLM_PMT_AMT ≥ 0 .

H_SNFPMT: SNF Medicare Payments is the total Medicare payments in the SNF setting for the calendar year.

H_SNFBPT: SNF Medicare Covered Days is the count of Medicare covered days in the SNF setting for the calendar year. This variable equals the sum of the CLM_UTLZTN_DAY_CNT variables on the source claims.

H_SNFSTY: SNF Stays is the count of SNF stays (unique admissions, which may span more than one facility) for the calendar year. A SNF stay is defined as a set of one or more consecutive SNF claims where the beneficiary is only discharged on the most recent claim in the set.

H_SNFBPT: Skilled Nursing Facility Beneficiary Payments is the sum of Medicare coinsurance and deductible payments in the SNF setting for the calendar year. The total beneficiary payment is calculated as the sum of the beneficiary deductible amount and coinsurance amount (variables NCH_BENE_IP_DDCTBL_AMT and NCH_BENE_PTA_COINSRNC_LBLTY_AM) for all SNF claims where the CLM_PMT_AMT ≥ 0 .

H_HOSPMT: Hospice Medicare Payments is the total Medicare payments in the hospice (HOS) setting for the calendar year.

H_HOSSTY: Hospice Stays is the count of stays (unique admissions, which may span more than one facility) in the HOS setting for the calendar year. A HOS stay is defined as a set of one or more consecutive hospice claims where the beneficiary is only discharged on the most recent claim in the set.

H_HOSDAY: Hospice Medicare Covered Days is the count of Medicare covered days in the HOS setting for the calendar year. This variable equals the sum of the CLM_UTLZTN_DAY_CNT variables on the source claims.

H_HHPMT: Home Health Medicare Payments is the total Medicare payments in the home health (HH) setting for the calendar year.

H_HHVIS: Home Health Visits is the count of HH visits for the calendar year.

H_HOPPMT: Hospital Outpatient Medicare Payments is the total Medicare payments in the hospital outpatient (HOP) setting for the calendar year.

H_HOPVIS: Hospital Outpatient Visits is the count of unique revenue center dates (as a proxy for visits) in the HOP setting for the calendar year.

H_HOP_ER: Hospital Outpatient Emergency Room Visits is the count of unique emergency department revenue center dates (as a proxy for an ED visit) in the HOP claims for the calendar year. Revenue center codes indicating emergency room use are 0450, 0451, 0452, 0456, or 0459.

H_HOPBPT: Hospital Outpatient Beneficiary Payments is the sum of Medicare coinsurance and deductible payments in the HOP setting for the calendar year. The total beneficiary payment is calculated as the sum of the beneficiary deductible amount and coinsurance amount (variables REV_CNTR_CASH_DDCTBLE_AMT and REV_CNTR_COINSRNC_WGE_ADJSTD_C) for all HOP claims where the CLM_PMT_AMT ≥ 0 .

H_PB_DRG: Part B Drug Medicare Payments is the total Medicare payments for Part B drugs for the calendar year. Part B drug claims are a subset of the claims in the Part B Carrier and DME claims.

H_PB_DEV: Part B Drug Events is the count of events in the Part B drug setting for the calendar year. An event is defined as each line item that contains the relevant service.

H_BPTDRG: Part B Drug Beneficiary Payments is the sum of coinsurance and deductible payments for Part B drugs for the calendar year. The total beneficiary payments are calculated as the sum of `LINE_COINSRNC_AMT` + `LINE_BENE_PTB_DDCTBL_AMT` for the relevant lines. The Part B drug claims are identified by BETOS codes (CCW variable `BETOS_CD` with values of "D1G", "O1D", "O1E", "O1G", "I1E", or "I1F").

H_EMPMT: Evaluation and Management Medicare Payments is the total Medicare payments for the Part B evaluation and management (E&M) services for a given year. E&M claims are a subset of the claims in the Part B Carrier and DME claims and a subset of physician claims.

H_EMEVT: E&M Events is the count of events for the Part B E&M services for the calendar year. An event is defined as each line item that contains the relevant service.

H_EMBPT: Evaluation and Management Beneficiary Payments is the sum of coinsurance and deductible payments for the Part B E&M services for the calendar year. The total beneficiary payments are calculated as the sum of `LINE_COINSRNC_AMT` and `LINE_BENE_PTB_DDCTBL_AMT` for the relevant lines. E&M claims are a subset of the claims in the Part B Carrier and DME data files and a subset of physician claims. The E&M claims are defined as those with a line BETOS code (`BETOS_CD`) where the first digit = "M" (but is not M1A or M1B – which are categorized as physician office care in this file – see `PHYS_MDCR_PMT`).

H_PHYPMPT: Part B Physician Medicare Payments is the total Medicare payments for the Part B physician office services (PHYS) for the calendar year. PHYS claims are a subset of the claims in the Part B Carrier and DME claims and a subset of physician evaluation and management claims (note that E&M are tabulated separately).

H_PHYEVT: Part B Physician Events is the count of events for Part B PHYS for the calendar year. An event is defined as each line item that contains the relevant service.

H_PHYBPT: Part B Physician Beneficiary Payments is the sum of coinsurance and deductible payments for the Part B PHYS for the calendar year. The total beneficiary payments are calculated as the sum of `LINE_COINSRNC_AMT` and `LINE_BENE_PTB_DDCTBL_AMT` for the relevant lines. The PHYS claims are defined as those with a line BETOS code (`BETOS_CD`) where the first three digits = M1A or M1B (the remainder of physician services which occur in different settings appear in `EM_MDCR_PMT`).

H_OPRPMT: Other Procedures Medicare Payments is the total Medicare payments for services considered Part B other procedures (i.e., not anesthesia or dialysis) for the calendar year. Claims for other procedures are a subset of the claims and a subset of procedures in the Part B Carrier claims.

H_OPREVT: Other Procedures Events is the count of events for Part B other procedures for the calendar year. An event is defined as each line item that contains the relevant service. Claims for other procedures are a subset of the claims in the Part B Carrier claims.

H_OPRBPT: Other Procedures Beneficiary Payments is the sum of coinsurance and deductible payments for services considered Part B other procedures for the calendar year. The total beneficiary payments are calculated as the sum of `LINE_COINSRNC_AMT` and `LINE_BENE_PTB_DDCTBL_AMT` for the relevant lines. Claims for other procedures are a subset of the claims in the Part B Carrier data file. These other procedure

claims are defined as those with a line BETOS code (BETOS_CD) where the first 2 digits are ("P1", "P2", "P3", "P4", "P5", "P6", "P7", or "P8").

H_DMEPMT: Durable Medical Equipment Medicare Payments is the total Medicare payments for Part B durable medical equipment (DME) for the calendar year. Claims for DME are a subset of the claims in the Part B Carrier and DME claims.

H_DMEEVT: Durable Medical Equipment Events is the count of events in the Part B DME for the calendar year. An event is defined as each line item that contains the relevant service.

H_DMEBPT: Durable Medical Equipment Beneficiary Payments is the total Medicare payments for Part B DME for the calendar year.

H_PB_OTH: Other Part B Carrier Medicare Payments is the total Medicare payments from Part B Carrier and DME claims which appear in specific settings for the calendar year. Claims for other carrier/DME claims are a subset of the claims in the Part B Carrier and DME claims. Types of services which may have been summarized in this other carrier category (OTHC) include ambulance, chiropractor, chemotherapy, vision, hearing and speech services, etc.

H_PB_OEV: Other Part B Carrier Events is the count of events in the Part B other setting for the calendar year, which includes Part B Carrier and DME claims which appear in specific settings for the year. An event is defined as each line item that contains the relevant service.

H_BPTOTH: Other Part B Carrier Beneficiary Payments is the sum of coinsurance and deductible payments from Part B Carrier and DME claims for the calendar year, which appear in settings other than the 10 specific categories in this segment. The total beneficiary payments are calculated as the sum of LINE_COINSRNC_AMT and LINE_BENE_PTB_DDCTBL_AMT for the relevant lines.

H_PTDPMT: Part D Medicare Payments is the dollar amount that the Part D plan covered for all covered drugs for the calendar year. The variable is calculated as the sum of the plan payments for covered Prescription Drug Events (PDEs) (CVRD_D_PLAN_PD_AMT) and the low-income cost sharing subsidy amount (LICS_AMT) during the year.

H_PTDEVT: Part D Events is the count of events for Part D drugs for the calendar year (i.e., a unique count of the PDE_IDs). An event is a dispensed (filled) drug prescription that appears on the source PDE claims.

H_PTDBPT: Part D Beneficiary Payments is the dollar amount that the beneficiary paid for all PDEs for the calendar year, without being reimbursed by a third party. The amount includes all copayments, coinsurance, deductible, or other patient payment amounts, and comes directly from the source PDEs.

H_PTDTOT: Part D Total Prescription Costs is the gross drug cost (TOT_RX_CST_AMT on the source claims) of all Part D drugs for the calendar year. This value includes the ingredient cost, dispensing fee, sales tax (if applicable), and vaccine administration fee.

H_ASCEVT: Ambulatory Surgery Center Events is the count of events in the Part B ambulatory surgery center (ASC) setting for the calendar year. An event is defined as each line item that contains an ambulatory surgery center service.

H_ASCBPT: Ambulatory Surgery Center Beneficiary Payments is the sum of coinsurance and deductible payments in the Part B ASC setting for the calendar year. The total beneficiary payment is calculated as the sum of the LINE_COINSRNC_AMT and LINE_BENE_PTB_DDCTBL_AMT for all relevant lines. ASC claims are a subset of the claims in the Part B Carrier data file. The ASC claims are identified by the claim lines where the LINE_CMS_TYPE_SRVC_CD = "F".

H_ANEPMT: Anesthesia Medicare Payments is the total Medicare payments for Part B anesthesia services (ANES) for the calendar year. Anesthesia claims are a subset of the claims and a subset of procedures in the Part B Carrier claims.

H_ANEVT: Anesthesia Events is the count of events for Part B ANES for the calendar year. An event is defined as each line item that contains the relevant service.

H_ANEBPT: Anesthesia Beneficiary Payments is the sum of coinsurance and deductible payments for Part B ANES for the calendar year. The total beneficiary payments are calculated as the sum of LINE_COINSRNC_AMT and LINE_BENE_PTB_DDCTBL_AMT for the relevant lines. ANES claims are a subset of the claims and a subset of procedures in the Part B Carrier data file. ANES claims are defined as those with a line BETOS code (BETOS_CD) where the first 2 digits = "P0" and the CARR_LINE_MTUS_CD = "2".

H_DIAPMT: Dialysis Medicare Payments is the total Medicare payments for Part B dialysis services (primarily the professional component since treatments are covered in hospital outpatient) for the calendar year. Dialysis claims are a subset of the claims and a subset of procedures in the Part B Carrier claims.

H_DIAEVT: Dialysis Events is the total Medicare payments for Part B dialysis services for the calendar year. An event is defined as each line item that contains the relevant service.

H_DIABPT: Dialysis Beneficiary Payments is the total Medicare payments for Part B dialysis services for the calendar year. The total beneficiary payments are calculated as the sum of LINE_COINSRNC_AMT and LINE_BENE_PTB_DDCTBL_AMT for the relevant lines.

H_IMGPMPT: Imaging Medicare Payments is the total Medicare payments for imaging services (IMG) for the calendar year. Claims for imaging procedures are a subset of the claims and a subset of procedures in the Part B Carrier and DME claims.

H_IMG EVT: Imaging Events is the count of events for IMG for the calendar year. An event is defined as each line item that contains the relevant service.

H_IMG BPT: Imaging Beneficiary Payments is the sum of coinsurance and deductible payments for IMG for the calendar year. The total beneficiary payments are calculated as the sum of LINE_COINSRNC_AMT and LINE_BENE_PTB_DDCTBL_AMT for the relevant lines. These IMG claims are defined as those with a line BETOS code (BETOS_CD) where the first digit = I (except for "I1E", or "I1F" – which are considered Part B drugs).

H_TSTPMT: Tests Medicare Payments is the total Medicare payments for Part B tests for the calendar year. Claims for tests are a subset of the claims in the Part B Carrier claims.

H_TSTEVT: Tests Events is the count of events for Part B tests for the calendar year. An event is defined as each line item that contains the relevant service. Claims for tests are a subset of the claims in the Part B Carrier claims.

H_TSTBPT: Tests Beneficiary Payments is the sum of coinsurance and deductible payments for Part B tests for the calendar year.

H_PTDFIL: Part D prescribing events (PDE) consist of highly variable days' supply of the medication. This derived variable creates a standard 30 days' supply of a filled Part D prescription and counts this as a "fill." The Part D fill count does not indicate the number of different drugs the person is using, only the total months covered by a medication (e.g., if a patient is receiving a full year supply of a medication, whether this occurs in

one transaction or 12 monthly transactions, the fill count = 12; if the patient is taking three such medications, the fill count = 36).

H_READMT: Acute Inpatient Hospital Readmissions is the count of hospital readmissions in the acute inpatient setting for the calendar year. The original admission must have been in the year of the data file, but it is possible for the readmission claim to have occurred in January of the following year. A beneficiary is considered to be readmitted when they have an acute inpatient stay with a discharge status that is not expired or left against medical advice within 30 days of a previous acute inpatient stay with a discharge status that is also not expired or left against medical advice.

10.3.3.3 Special Notes

For easier comparison of groups of beneficiaries by the number and cost of medical services they have received, the Administrative Utilization Summary includes a summary of all Medicare bills and claims for calendar year 2022, as received and processed by CMS through December 2023 for the 2022 benefit year.

The utilization summary represents services rendered and reimbursed under Medicare FFS in the calendar year 2022. If a beneficiary used no Medicare services at all or was a member of a coordinated or managed care plan that does not submit claims to a fiscal intermediary or carrier, all program payment summary variables will be null. If the beneficiary used no services of a particular type (e.g., inpatient hospitalization), the variables relating to those benefits will be null.

For additional information on administrative data items, please see the Master Beneficiary Summary – Cost and Use Segment Data Dictionary Codebook: <https://www.cwdata.org/web/guest/data-dictionaries>.

10.3.4 Assistance (ASSIST)

10.3.4.1 Core Content

The Assistance segment contains information on each person identified as helping the beneficiary with ADLs or IADLs, including the helper's age, relationship to the beneficiary, and the types of assistance that the beneficiary receives (e.g., assistance with dressing, shopping, eating) from each identified helper. The number of records in the Assistance segment reflects the number of persons identified as having assisted the beneficiary with one or more ADL or IADLs. Therefore, it is possible to have one, several, or no helper records per beneficiary.

10.3.4.2 Variable Definitions

Please see the Codebook for information regarding variables in this segment.

10.3.4.3 Special Notes

When a beneficiary has more than one helper, the variable HLPRMOST identifies which helper provides the beneficiary with the most help with daily activities. This variable contains missing values for helpers who were a beneficiary's only helper. If a beneficiary with multiple helpers has not indicated which helper provides the most help, then this variable contains missing data for each of the beneficiary's helpers.

Most ADL and IADL questions are asked in the HFQ section in the fall round. However, there is one variable, HLPRUSGO (the person who usually accompanies the beneficiary to their provider's office), that is asked in the winter round in the USQ section.

10.3.5 Chronic Conditions (CHRNCOND)

10.3.5.1 Core Content

The Chronic Conditions segment contains information on whether the beneficiary has a series of chronic and other diagnosed medical conditions such as cancer, high blood pressure, and depression. If the respondent reports that the beneficiary has the condition, a series of follow-up questions is asked.

10.3.5.2 Variable Definitions

D_OCDTYP: This variable indicates type of diabetes and is derived from OCDTYPE and DIAPRGNT. The OCDTYPE categories for "Pre-diabetes" and "Borderline" diabetes are combined into one category for D_OCDTYP. Female beneficiaries who answer "Yes" for DIAPRGNT, which is not released, are coded as "Gestational diabetes" for D_OCDTYP, unless they indicate for OCDTYPE that they have Type 1 diabetes.

LOSTURIN: "More than once a week" is coded if the beneficiary cannot control urination at all. Leaking urine, especially when the person laughs, strains or coughs, does not qualify as incontinence.

10.3.5.3 Special Notes

The HFQ and PVQ sections ask respondents whether they have ever had any of a series of illnesses or conditions in the fall round. Their responses are coded affirmatively if the beneficiary had at some time been diagnosed with the condition, even if the condition had been corrected by time or treatment. The condition must be reported by the respondent as diagnosed by a physician, not by the respondent. If the respondent is not sure about the definition of a condition, the interviewer offers no advice or information, but records the respondent's answer verbatim.

There are different versions of each illness/condition question depending on whether a respondent is in the Incoming Panel sample or Continuing sample. Incoming Panel sample respondents are asked if a doctor ever told them that they had a specific condition (e.g., hypertension). If the answer is "Yes", then the Incoming Panel respondent is asked if the doctor had told them in the past year that they had the condition.

For illnesses or conditions that cannot change after diagnosis (e.g., Alzheimer's), once an affirmative response is given, respondents are not asked again. However, if a negative response is given, respondents are asked annually thereafter if they had that specific illness or condition in the past year.

For illnesses or conditions that can change after diagnosis or can be reoccurring, such as high blood pressure, respondents are asked annually thereafter if they were diagnosed with that illness or condition in the past year, irrespective of prior responses. All data for a beneficiary from the current survey year and all previous years are used to determine whether the beneficiary has ever been told by a doctor that they had a condition. The CHRNCOND segment includes variables that indicate whether a beneficiary ever had specific conditions.

The "other specify" questions EMOS and EVROS are back coded as necessary into the "reason for Medicare eligibility" and "type of cancer" response options, respectively, but the verbatim text is not released.

In 2022, the code list at HYPEDRINK, which asks whether the beneficiary has cut down on drinking alcoholic beverages because of their high blood pressure, was updated to include a "Not applicable" response for beneficiaries who do not drink alcoholic beverages.

Prior to 2022, items OCCHOLES and YRCHOLES, which collect whether a doctor or health professional has told the beneficiary they have high cholesterol ever or in the past year, respectively, were erroneously administered to both Baseline and Continuing cases. In Fall 2022, the universe of these items was corrected to administer OCCHOLES only to Baseline cases and YRCHOLES only to Continuing cases. In the CHRNCOND

segment, these items are combined into a single variable, OCCHOLE, which indicates whether the beneficiary been diagnosed with high cholesterol.

10.3.6 Chronic Conditions Flags (CHRNCDL)

10.3.6.1 Core Content

The Chronic Conditions Flags segment contains chronic and other disabling conditions flags from administrative FFS records from the CCW. The CCW summarizes beneficiaries' FFS claims for the calendar year and indicates whether a claim for a particular condition met criteria for inclusion. This segment also provides the first year the beneficiary met the criteria for having that chronic condition. Variables are included for those conditions related to the self-reported information included in the MCBS instrument and are not inclusive of all chronic and disabling conditions available.

10.3.6.2 Variable Definitions

The end of year indicator flags require beneficiaries to satisfy both claims criteria (a minimum number/type of claims that have the proper diagnosis codes and occurred within a specified time period) and coverage criteria (FFS Part A and Part B coverage during the entire specified time period). Each flag is also created using details about the specific condition that must be met for inclusion.

Indicators have the following values:

- 0 = Beneficiary did not meet claims criteria or have sufficient FFS coverage
- 1 = Beneficiary met claims criteria but did not have sufficient FFS coverage
- 2 = Beneficiary did not meet claims criteria but had sufficient FFS coverage
- 3 = Beneficiary met claims criteria and had sufficient FFS coverage

The ever indicator variables for the conditions show the date when the beneficiary first met the criteria for the chronic or disabling condition. The variable is missing for beneficiaries that have never had the condition. The earliest possible date for anyone is January 1, 2016. If the beneficiary became eligible for Medicare after that, the earliest possible date will be some time after their coverage start date.

10.3.6.3 Special Notes

The end of year indicator flags criteria was developed after reviewing validated algorithms from the research literature and criteria used by other federal data sources. Please visit the CCW website for more detailed information on the criteria: <https://www.ccwdata.org/web/guest/condition-categories>.

Beginning with the 2022 Survey File, the CHRNCDL flags only reflect ICD-10 codes, rather than ICD-9 and ICD-10 codes as in previous years. As such, diagnoses made before 2016 are no longer reflected in the flags.

In 2022, the ALZHDMTA (claim for Alzheimer's disease/dementia this benefit year) and ALZHDMTE (1st date ever claim for Alzheimer's disease/dementia) variables were dropped due to changes in the source data. Additionally, the CHF (claim for congestive heart failure (CHF) this benefit year) and CHFE (1st date ever claim for CHF) variables were renamed to HF and HFE, respectively, to reflect changes in how these flags are constructed in the source data. For HF, the flag now captures "heart failure (HF) and non-ischemic heart disease" instead of just CHF.

10.3.7 Chronic Pain (CHRNPAIN)

10.3.7.1 Core Content

The Chronic Pain segment contains data on beneficiaries' experiences with chronic pain and chronic pain management techniques collected in the CPQ section administered the summer following the year of interest. The CPQ collects information related to frequency and severity of chronic pain, location of chronic pain (e.g., hips, knees, or feet), and use of pain management techniques (e.g., massage).

10.3.7.2 Variable Definitions

Please see the Codebook for information regarding variables in this segment.

10.3.7.3 Special Notes

The CPQ uses a three-month reference period; thus, the items administered in Summer 2023 (Round 96) asked beneficiaries about pain experienced in 2023. However, because the CPQ is administered to beneficiaries who were ever enrolled in Medicare in 2022 and are still enrolled in 2023, the CPQ data are released with the 2022 Survey File.

Special non-response adjustment weights are included in the file to account for survey non-response from the fall to the summer data collection period. This includes Survey File ever enrolled and continuously enrolled weights, as well as Cost Supplement ever enrolled weights. Note that counts of cases with positive Topical weights may vary within the data year and may change across years due to response rates, sample sizes, and fielding methods. The Topical weights account for these changes. Please see section 9.4.2 for information on using Topical weights.

10.3.8 Cognitive Measures (COGNFUNC)

10.3.8.1 Core Content

The Cognitive Measures segment contains data on the beneficiary's cognitive abilities collected in the CMQ section administered in the fall rounds. The CMQ contains four cognitive measures, including backwards counting, date naming, object naming, and president/vice president naming.

10.3.8.2 Variable Definitions

Please see the Codebook for information regarding variables in this segment.

10.3.8.3 Special Notes

N/A

10.3.9 COVID-19 Experiences (COVIDEXP)

10.3.9.1 Core Content

The COVID-19 Experiences segment contains information collected in the CVQ section during the fall round, and it includes data on COVID-19 vaccination, testing, diagnosis, symptoms, and prevention, as well as persistent symptoms of long COVID-19.

10.3.9.2 Variable Definitions

Please see the Codebook for information regarding variables in this segment.

10.3.9.3 Special Notes

COVID-19 vaccine data are brought forward from the 2021 Survey File, but no other COVIDEXP data are combined across interview rounds. The vaccine doses included in COVIDEXP are those that were reported at or before the beneficiary's Fall 2022 interview. For data on COVID-19 experiences and vaccination reported at the Winter 2023 and Summer 2023 Community interviews, please see the COVIDTOP segment.

In Fall 2022, question text for SWABRSLT, which collects whether a COVID-19 test found the beneficiary had COVID-19, was modified to instruct beneficiaries to answer in the affirmative if any COVID-19 tests taken during the reference period were positive. To reflect the corresponding change in the universe of respondents at this item, the variable was renamed COVRSLT.

Also in Fall 2022, three items were modified to account for the varying reference periods between Baseline and Continuing cases. Each of the three items was split into two new variables to reflect an "ever" reference period for Baseline cases and a "date of the last interview" for Continuing cases. Item SUSPECT was replaced with EVRSUS (Baseline) and COVSUS (Continuing), item COVIDEV was replaced with EVRCVTLD (Baseline) and COVTOLD (Continuing), and item COVSWAB was replaced with EVRCVTST (Baseline) and COVTEST (Continuing). In the COVIDEXP segment, data from these three pairs of items are combined into the variables EVRSUS, EVRCVTLD, and EVRCVTST, respectively.

The "other specify" question CVDOTHER is back coded as necessary into the "why beneficiary didn't seek medical care for coronavirus" response options, but the verbatim text is not released.

In Fall 2022, the CVQ items that collect persistent symptoms of COVID-19 (SMPTFATG, SMPTHEAD, SMPTHRT, SMPTACHE, SMPTCOGH, SMPTDIZZ, SMPTANX, SMPTOTH) were deleted and replaced with the NHIS item collecting symptoms of long COVID-19 (LONGCVD).

10.3.10 COVID-19 Topical (COVIDTOP)

10.3.10.1 Core Content

The Community COVID-19 Topical segment contains information collected in the CVQ section during the Winter 2023 and Summer 2023 rounds. This data includes information about COVID-19 vaccination, testing, diagnosis, symptoms, and treatment, as well as persistent symptoms of long COVID-19.

10.3.10.2 Variable Definitions

Please see the Codebook for information regarding variables in this segment.

10.3.10.3 Special Notes

There is one row for every Community beneficiary who was included in the Summer 2023 population. The vaccine receipt variables refer to a vaccine dose that was first recorded in either the Winter 2023 or Summer 2023 interview; if a beneficiary has multiple doses recorded during these two interviews, information about the most recent vaccine dose is given by these variables.

Several "other specify" variables are back coded as necessary into response options, but the verbatim text is not released. Back coded "other specify" variables include VACSITOS (where the beneficiary went for their

COVID-19 vaccine), VACNMEOS (which COVID-19 vaccine did the beneficiary get for dose), and NOVACOS (why the beneficiary is not likely to receive a vaccine).

In Winter 2023, the question text and response options for VACNME were updated to include the COVID-19 vaccine Novavax. In COVIDTOP, this item is released under the variable name D_VACNME.

Special non-response adjustment weights are included in the file to account for survey non-response from the fall to the summer data collection period. This includes Survey File ever enrolled and continuously enrolled weights, as well as Cost Supplement ever enrolled weights. Note that counts of cases with positive Topical weights may vary within the data year and may change across years due to response rates, sample sizes, and fielding methods. The Topical weights account for these changes. Please see section 9.4.2 for information on using Topical weights.

10.3.11 Demographics (DEMO)

10.3.11.1 Core Content

The Demographics segment contains demographic information collected in the survey as well as demographic information from Medicare Administrative enrollment data and constructed items of interest.

10.3.11.2 Variable Definitions

ADI: The Area Deprivation Index (ADI) is an indicator of the socioeconomic deprivation of geographic areas and is intended for use in evaluating the relationship between socioeconomic factors and health. This index was originally developed using 17 markers of socioeconomic status from the 1990 Census data. The ADI dataset used in this data release was developed by Amy Kind, MD, PhD and her research team at the University of Wisconsin using the same indicators and 2020 Census block group-level data from the American Community Survey (ACS). This dataset contains national percentile rankings at the block group level from 1 to 100 as well as state decile rankings from 1 to 10. Raw ADI values are used to determine percentile and decile rankings. ADI values in the first percentile are the least disadvantaged, and those in the hundredth are the most disadvantaged.²⁹

The MCBS includes two ADI values for each beneficiary, a national-level percentile (ADINATNL) and a state-level decile (ADISTATE). Both rankings are based on the Census block group for the beneficiary's primary residence address. Beneficiaries have a value for each of these variables if their Census block group is found on the ADI dataset. Excluding the exiting 2017 panel cases, there was a 95.7 percent match rate for cases matched to the ADI dataset.

H_DOB, H_DOD, H_AGE, and D_STRAT: These four variables are related to the beneficiary's age. The "legal" dates of birth and death from Medicare and the Social Security Administration records are recorded as H_DOB and H_DOD, respectively. The variable H_AGE represents the "legal" age as of December 31, 2022, adjusted for date of death, if present. The variable D_STRAT groups the beneficiaries by various age categories using H_AGE. The date of birth, as reported during the Baseline interview, is recorded in DEMO (D_DOB).

D_DOB: When the complete date of birth is entered (D_DOB) in the MCBS instrument, the CAPI questionnaire automatically calculates the person's age, which is then verified with the respondent. Despite this validation, the date of birth given by the respondent (D_DOB) does not always agree with the date of birth per CMS records (H_DOB). In these cases, the beneficiary is asked again in the next interview to provide a date of birth. Some recording errors are identified this way, but in most cases, beneficiaries provide the same date of

²⁹ "2020 Area Deprivation Index v3.2," University of Wisconsin School of Medicine and Public Health, <https://www.neighborhoodatlas.medicine.wisc.edu/>.

birth both times they are asked. In some cases, proxies indicate that no one is exactly sure of the correct date of birth. In general, it is recommended that the variable (H_DOB) be used for analyses, since the CMS date of birth is used to select and stratify the sample.

D_DOD: Date of death provided by proxy respondents. In general, it is recommended that both the survey-reported (D_DOD) and administrative (H_DOD) variables be used for analyses.

D_RACE2: Race categories are self-reported by the respondent. Categories are not suggested by the interviewer, nor does the interviewer try to explain or define any of the groups. Ethnic groups such as Irish or Cuban are not recorded.

H_CENSUS: The Census division is performed through internal edits by matching the survey respondent's SSA State code to the appropriate Census region. The Census divisions are as follows:

- New England – Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont
- Middle Atlantic – New Jersey, New York, Pennsylvania
- South Atlantic – Delaware, District of Columbia, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, West Virginia
- East North Central – Illinois, Indiana, Michigan, Ohio, Wisconsin
- West North Central – Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota
- East South Central – Alabama, Kentucky, Mississippi, Tennessee
- West South Central – Arkansas, Louisiana, Oklahoma, Texas
- Mountain – Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, Wyoming
- Pacific – Alaska, California, Hawaii, Oregon, Washington

Urban/rural status variables: H_RUCA indicates overall urban/rural status. H_RUCA1 and H_RUCA2 indicate the primary and secondary RUCA codes, respectively (see Exhibits 10.3.11.2a and 10.3.11.2b). This classification scheme provides an alternative to county-based systems for situations where more detailed geographic analysis is feasible. It identifies areas of emerging urban influence and areas where urban-rural classifications overlap, thus providing an exhaustive system of statistical areas for the country.

The ten whole numbers shown in Exhibit 10.3.11.2a below refer to the primary, or single largest, commuting share. Metropolitan cores (code 1) are defined as Census tract equivalents of urbanized areas. Micropolitan and small town cores (codes 4 and 7, respectively) are tract equivalents of urban clusters. Tracts are included in urban cores if more than 30 percent of their population is in the urbanized area or urban cluster.

High commuting (codes 2, 5, and 8) means that the largest commuting share is at least 30 percent to a metropolitan, micropolitan, or small town core. Many micropolitan and small town cores themselves (and even a few metropolitan cores) have high enough out-commuting to other cores to be coded 2, 5, or 8; typically, these areas are not job centers themselves but serve as bedroom communities for a nearby larger city. Low commuting (codes 3, 6, and 9) refers to cases where the single largest flow is to a core but is less than 30 percent. These codes identify "influence areas" of metro, micropolitan, and small town cores, respectively, and are similar in concept to the "nonmetropolitan adjacent" codes found in other Economic Research Service (ERS) classification schemes ([Rural-Urban Continuum Codes](#), [Urban Influence Codes](#)). The last of the general classification codes (10) identifies rural tracts where the primary flow is local or to another rural tract.

Exhibit 10.3.11.2a: Primary RUCA (H_RUCA1) Codes, 2010

Code	Classification description
1	Metropolitan area core: primary flow within an urbanized area (UA)
2	Metropolitan area high commuting: primary flow 30% or more to a UA
3	Metropolitan area low commuting: primary flow 10% to 30% to a UA
4	Micropolitan area core: primary flow within an urban cluster (UC) of 10,000 to 49,999 (large UC)
5	Micropolitan high commuting: primary flow 30% or more to a large UC
6	Micropolitan low commuting: primary flow 10% to 30% to a large UC
7	Small town core: primary flow within an urban cluster of 2,500 to 9,999 (small UC)
8	Small town high commuting: primary flow 30% or more to a small UC
9	Small town low commuting: primary flow 10% to 30% to a small UC
10	Rural areas: primary flow to a tract outside a UA or UC
99	Not coded: Census tract has zero population and no rural-urban identifier information

These ten codes offer a relatively straightforward and complete delineation of metropolitan and nonmetropolitan areas based on the size and direction of primary commuting flows. However, secondary flows may indicate other connections among rural and urban places. Thus, the primary RUCA codes are further subdivided to identify areas where classifications overlap, based on the size and direction of the secondary, or second largest, commuting flow (see Exhibit 10.3.10.2b). For example, 1.1 and 2.1 codes identify areas where the primary flow is within or to a metropolitan core, but another 30 percent or more commute to a larger metropolitan core. Similarly, 10.1, 10.2, and 10.3 identify rural tracts for which the primary commuting share is local, but more than 30 percent also commute to a nearby metropolitan, micropolitan, or small town core, respectively.

Exhibit 10.3.11.2b: Secondary RUCA (H_RUCA2) Codes, 2010

Code	Classification description
1 Metropolitan area core: primary flow within an urbanized area (UA)	
1.0	No additional code
1.1	Secondary flow 30% to 50% to a larger UA
2 Metropolitan area high commuting: primary flow 30% or more to a UA	
2.0	No additional code
2.1	Secondary flow 30% to 50% to a larger UA
3 Metropolitan area low commuting: primary flow 10% to 30% to a UA	
3.0	No additional code
4 Micropolitan area core: primary flow within an urban cluster (UC) of 10,000 to 49,999 (large UC)	
4.0	No additional code
4.1	Secondary flow 30% to 50% to a UA
5 Micropolitan high commuting: primary flow 30% or more to a large UC	
5.0	No additional code
5.1	Secondary flow 30% to 50% to a UA
6 Micropolitan low commuting: primary flow 10% to 30% to a large UC	
6.0	No additional code
7 Small town core: primary flow within an urban cluster of 2,500 to 9,999 (small UC)	
7.0	No additional code
7.1	Secondary flow 30% to 50% to a UA
7.2	Secondary flow 30% to 50% to a large UC
8 Small town high commuting: primary flow 30% or more to a small UC	
8.0	No additional code
8.1	Secondary flow 30% to 50% to a UA
8.2	Secondary flow 30% to 50% to a large UC
9 Small town low commuting: primary flow 10% to 30% to a small UC	
9.0	No additional code
10 Rural areas: primary flow to a tract outside a UA or UC	
10.0	No additional code
10.1	Secondary flow 30% to 50% to a UA
10.2	Secondary flow 30% to 50% to a large UC
10.3	Secondary flow 30% to 50% to a small UC
99 Not coded: Census tract has zero population and no rural-urban identifier information	

INCOME: Income represents the best source or estimate of income during the year of interest. Data gathered in fall and summer interviews represent the most detailed data and are used when available. For individuals who did not complete the fall interview (that is, Continuing Panel people unavailable for their fall interview), the most recent information available is used. It should be noted that the variable INCOME includes income from all sources, such as pension, Social Security, and retirement benefits, for the beneficiary and spouse/partner. In some cases, the respondent will not or cannot provide specific information but did say the income is above or below \$25,000.

INT_TYPE: Provides the source for a beneficiary's residence status at the time of interview, and the types of interviews conducted with C = Community, F = Facility, and B = Both. INT_TYPE is defined as:

- C = respondent only resided in the community and only completed Community-administered survey instruments in each round
- F = respondent only resided in a facility and only completed Facility-administered survey instruments in each round
- B = respondents completed instruments in both settings across the rounds

INT_TYPE is created following the rules below:

- Beneficiaries are assigned an INT_TYPE if they completed or partially completed an interview in at least one round in 2022. INT_TYPE is also calculated for beneficiaries who completed an interview, but died or lost entitlement during the data year.
- Missing INT_TYPES – There are currently 30 beneficiaries with “complete” dispositions which cannot have their INT_TYPE/residence location calculated for them. These are individuals that appear to have died in early 2022 and did not have any completed/partially completed questionnaire data for 2022. These individuals have ever enrolled weights, but do not have completed interviews.

Note that in each data year, some differences by segment will exist (i.e., data may reflect a prior or future calendar year due to the specific questionnaire and reference period used to collect the information).

INT_TYPE is only constructed using survey-reported data for the benefit year and is not edited to account for data collected in a future or prior data year.

INT_TYPE is calculated on the benefit year, but data segments may reflect a prior or future calendar year due to the specific questionnaire and reference period used to collect the information. That is, the segment data is collected prior to or after the benefit year designation of INT_TYPE.

For example, there may be beneficiaries living in facilities (INT_TYPE = F) that appear on the 2022 segments that include 2022 non-response adjustments: ACCSSMED, CHRNPAIN, COVIDTOP, FOODINS, INCASSET, MCREPLNQ, RXMED, TELEMED, and USCARE. The MCBS would expect these segments to only include beneficiaries with INT_TYPE = C or B because these segments contain data from survey-reported instruments only asked of beneficiaries that reside in the community. However, because the data for these segments is collected in 2023, beneficiaries may have moved from a facility in 2022 to the community in 2023 at the time these data segments were collected.

Alternatively, data may be pulled forward from a prior data collection year. For example, a beneficiary in 2021 that answered affirmative to the question, “Have you ever had a hysterectomy?”, a survey item that is asked of beneficiaries in the Community Questionnaire, will have that answer pulled forward to the 2022 data segment even if the beneficiary currently resides in a facility in 2022, and thus they would show an INT_TYPE = F. INT_TYPE is only constructed using survey-reported data for the benefit year and is not edited to account for data collected in a future or prior data year.

IPR: Indicates the income-to-poverty ratio (IPR). The Census Bureau determines who is “poor” by comparing an individual or household's income to a set of dollar-value thresholds that are intended to represent the amount of income needed to meet basic needs and are adjusted for family size and composition. A family will be designated as “poor” or “not poor” depending on whether their income is at or below or above this set threshold in a given year. In addition, the Census Bureau provides another way to describe a person's economic well-being by gauging how close to or far from the poverty threshold a family's income rests using an IPR. IPRs, income divided by the appropriate poverty threshold, are used to normalize incomes across family types and provide context for a better understanding of the depth of poverty (or lack thereof) of a family. The IPR is a useful analytic tool that can help MCBS users to easily identify the percentage of Medicare

beneficiaries living in deep poverty, below poverty, or “near” poverty (usually defined as less than 125 percent of the poverty level); or how health care access and use may differ across different thresholds of interest. Note that the MCBS IPR is calculated only for household sizes of one (beneficiary living alone or in a facility) or two (beneficiary living with a spouse/partner only) as the Income and Asset information is collected only from the beneficiary and the beneficiary’s spouse/partner. Medicare beneficiaries have slightly different poverty level indices used for program eligibility. The IPR uses the Medicare poverty thresholds for calculation but can be unformatted to create other thresholds.

PANEL: Indicates the year of the beneficiary’s Baseline interview.

10.3.11.3 Special Notes

The Demographics segment contains all demographic data from both the survey and from CMS administrative records.

The Department of Veterans Affairs (VA) disability rating collected at SPVARATE is a percentage and is expressed in multiples of ten; it refers to disabilities that are officially recognized by the government as service-related. If the VA finds that a Veteran has multiple disabilities, the VA uses a Combined Ratings Table to calculate a combined disability rating (see <https://www.benefits.va.gov/compensation/rates-index.asp#combined>).

The data at SURVIVE contains information about beneficiaries who were continuously enrolled in Medicare from January 1 up to and including their fall round interview. The “other specify” question WHTLNGOS is back coded as necessary into the “languages spoken at home” response options, but the verbatim text is not released. Similarly, the “other specify” questions HISPDTOS, RACEASOS, and RACEPIOS are also back coded as necessary into the “Other Hispanic/Latino/Spanish”, “Other Asian”, and “Other Pacific Islander” response options respectively, but the verbatim text is not released.

In Fall 2022, the administration of item ROSTSEX, which collects the beneficiary’s sex and was only read to the respondent if “sex is not obvious,” was modified such that it is read aloud to all respondents and the beneficiary’s sex is no longer preloaded from administrative data.

10.3.12 Diabetes (DIABETES)

10.3.12.1 Core Content

The Diabetes segment includes survey responses related to diabetes management. Only beneficiaries living in the community who indicated that they had ever been told they have non-gestational diabetes (variable D_OCDTYP in the Chronic Condition segment) are included in the Diabetes segment. This segment includes beneficiaries who indicated they had been diagnosed with any of these diabetic conditions: Type 1, Type 2, pre-diabetes/borderline diabetes, or other non-gestational type of diabetes.

10.3.12.2 Variable Definitions

Frequency of management variables: The Diabetes segment includes five pairs of items that describe the frequency of specific diabetes management behaviors. These behaviors are taking insulin, using an insulin pump, taking prescription or oral diabetes medications, testing blood glucose, and checking for foot sores. The frequency of each behavior is described by a pair of variables, with one set yielding the numeric frequency (variables D_INSFREQ, D_INSPMP, D_MEDFRQ, D_TSTFRQ, and D_SORFRQ, respectively). The other set of variables captures the corresponding frequency unit, with the exception of D_INSPMP (variables INSUUNIT, MEDSUNIT, TESTUNIT, and SOREUNIT, respectively).

10.3.12.3 Special Notes

The variables included in the Diabetes segment are centered on diabetes management. It should be noted there are other diabetes-related variables on other segments. For example, the Chronic Condition segment stores variables relevant to diabetes diagnoses (e.g., OCBETES and D_OCDTYP). Variables related to diabetes risk and screening (e.g., DIAEVERT, DIARCNT, DIAAWARE, DIARISK, and DIASIGNS) appear in the PREVCARE segment. The variable pertaining to diabetic retinopathy (ERETINOP) appears in the VISHEAR segment.

Items TESTDAY, TESTWEEK, TESTMNT, TESTYEAR collect how often a beneficiary who has reported diabetes tests their blood for sugar or glucose. In Fall 2022, these items were collapsed into one new item CONTGLUC to accommodate beneficiaries who use a continuous glucose monitoring system to check their blood sugar.

Last available in the 2018 Survey File LDS, the variables DIATRAN, DIANEURO, DIACIRCF, DIAULCER, and DIASKINC are available again beginning with the 2022 Survey File LDS. Please see the Codebook for additional information on these variables.

10.3.13 Facility Assessments (FACASMNT)

10.3.13.1 Core Content

CMS designed the MDS instrument to collect information regarding the health status and functioning of nursing home residents. The MDS is administered to anyone residing in a certified nursing home, regardless of payer. About half of MCBS beneficiaries living in a facility at the time of their interview live in certified nursing homes. For this reason, the MCBS Facility instrument has been designed to mirror parts of the MDS instrument.

10.3.13.2 Variable Definitions

D_HYST: Beneficiary ever had a hysterectomy. This variable is set to 1 if there was ever a "yes" response to the Facility variables EVERHYST or HYSTEREC. D_PNEU: Beneficiary ever had a pneumonia shot. This variable is set to 1 if there was ever a "yes" response to the Facility variable PNEUSHOT. Otherwise, it is set to the most recent value of Facility variable PNEUSHOT.

D_SMOKE: This variable is set to:

- 1 if there was ever a "yes" response to the facility variable EVRSMOKE, otherwise
- .D if there was ever a "don't know" response to the facility variable EVRSMOKE, otherwise
- The most recent value of Facility variable EVRSMOKE

10.3.13.3 Special Notes

Special Notes Regarding the Integration of MDS Data with FACASMNT

For beneficiaries for which the facility respondent reported a CCN, more than half of the variables in FACASMNT are skipped during data collection. The survey-reported data are later merged with MDS administrative data in data processing using the BASEID and CCN. Specifically, MDS data from the Nursing Home Comprehensive and Quarterly assessments are integrated with the FACASMNT data using the following data matching protocol:

- If there is an MDS record with an assessment date exactly matching the survey-reported assessment date, this MDS record is used.

- Otherwise, if there is an MDS assessment within 90 days of the survey-reported assessment date, the identified MDS record is used.
- If there is no survey-reported assessment date and there is an MDS assessment within 90 days of the survey reference date, this MDS record is used.
- If no MDS assessments meet these criteria, the survey record is unable to be matched to the administrative data, and the skipped variables are not populated during data processing and thus remain missing on FACASMNT.

FACASMNT data match outcomes for 2022 are detailed in Exhibit 10.3.13.3a.

Exhibit 10.3.13.3a: FACASMNT Administrative Data Match Outcomes

Match Type	Record Count
MDS record identified via an exact date match between the survey-reported assessment date and MDS assessment date	432
MDS record identified via a non-exact date match between the survey-reported assessment date and MDS assessment date	48
MDS record identified via the survey reference date	17
No match found	14

A flag variable, D_SOURCE, indicates whether the FACASMNT record has been populated for qualifying variables using the MDS.

Since the MCBS Facility Instrument has been designed to mirror the MDS, the MDS data used is mostly comparable to the survey-reported data, but there are minor differences in the handling of item non-response and missing data. As the MDS data is administrative, values of .R (refuse to answer) and .D (don't know the answer) are not possible for these records. The MDS administrative data uses a dash, "-", to signify a missing value, while the survey-reported data use a period, ".". Values of "-" have been converted to "." in the FACASMNT segment to maintain the same convention as survey-reported data, but "-" values remain intact in the MDS3 segment.

The FACASMNT variables that may be populated during data processing using MDS data are indicated in the table below.

Exhibit 10.3.13.3b: FACASMNT Variables Populated with Administrative Data

Variable Names			
AFIBDYS	CSNAMFAC	MENTDOWN	PHQSYCON
ALZHMR	CTBLADDC	MENTEPIL	PHQSYDEP
ANEMIA	CTBOWELC	MENTOTHN	PHQSYINT
ANXIETY	CVATIAST	MENTOTHO	PHQSYMOMV
APHASIA	DEHYD	MENTSUM	PHQSYSES
ARTHRIT	DELUS	MOLCANE	PHQSYSPL
ASTHCOPD	DEMENT	MOLPROS	PHQSYSUI
BPH	DEPRESS	MOLWCHR	PHQSYTEM
BRAININJ	DIABMRN	MOLWLKR	PHQSYTIR
BSAYSOT	DVTPEPTE	NUROBLAD	PSYCOTIC

Variable Names

BSELFAC	GERDULC	OBURPATH	PTSD
BSELFCA	HALLUC	ORTHHYPO	PVDPAD
BSELFILL	HARTFAIL	OSTEOP	QUADPLEG
BSNOEVAL	HCHEAID	OTHFRACT	RENLESRD
BSNOTOT	HCHECOND	PARAPLEG	RESPFAIL
BSOFTWAN	HCUNCOND	PARKNSON	SCHIZOPH
BSOTHACT	HCUNDOH	PFBATHNG	SCLEROS
BSOTHENV	HEIGHT	PFDRSSNG	SEIZEPIL
BSOTHILL	HEMIPLA	PFEATING	SOCIVITY
BSVERBOT	HIPFRACT	PFLOCOMO	SOCHEW
BSWDANGR	HUNTDIS	PFTOILET	SOCOUGH
BSWOTACT	HYPERKAL	PFTRNSFR	SODENT
CANCER	HYPETENS	PHQINTRO	SOGUMS
CATGLAUC	HYPONMIA	PHQSCORE	SOHOLD
CERPALSY	HYPRLIPI	PHQSFQAP	SOLOSS
CIRROSIS	INFHPPTS	PHQSFQCO	SOPAIN
COLCROHN	INFMDRO	PHQSFQDE	SOTEETH
COMATOSE	INFPNEU	PHQSFQIN	SOTISSUE
CORARTDS	INFSEPT	PHQSFQMO	THYROID
CSCURSEA	INFTBRC	PHQSFQSE	TOURETTE
CSDECIS	INFURNRY	PHQSFQSL	VISAPPL
CSINNH	INFWND	PHQSFQSU	VISION
CSLOCROM	MALNUTRI	PHQSFQTE	WEIGHT
CSMEMLT	MANICDEP	PHQSFQTI	
CSMEMST	MENTAUTI	PHQSYAPT	

What is the difference between the MDS3 and FACASMNT segments?

FACASMNT rows populated with MDS data can be linked to the corresponding MDS3 rows using the unique key BASEID, TRGT_DT, and A2300.

See the exhibit below for key differences between the segment sources, population, reference period, and unit of observation.

Exhibit 10.3.13.3c: Differences between FACASMNT and MDS3 Data

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

What is the difference between FACASMNT and similar Community segments?

Many of the variables on the FACASMNT segment are similar to variables available on the Survey File segments containing data from the Community interview. The exhibit below summarizes the topics that are available on FACASMNT that have similar content on a Community segment. However, in order to combine Community and Facility data together for analysis, some variables may need to be recoded to account for differences in response categories between Community and Facility variables. See Chapter 7 for more information about combining Community and Facility data.

For information on using flu shot data in analyses, please see PREVCARE.

Exhibit 10.3.13.3d: LDS Segments with Similar Topics for Community and Facility Interviews

Topic	Segments with Community Data	Segments with Facility Data	Segments with Data for All Beneficiaries
Health Status	GENHLTH FALLS CHRNCOND MENTHLTH OASIS	FACASMNT MDS3	
Preventive Care	PREVCARE	FACASMNT	
Functional Status & Assistance with Long-Term Care Needs	ASSIST NAGIDIS OASIS MOBILITY	FACASMNT MDS3	
Demographics and Socio-Economic Status	INCASSET		DEMO
Health Insurance Coverage			HISUMRY HITLINE ADMNUTLS

10.3.14 Facility Characteristics (FACCHAR)

10.3.14.1 Core Content

The Facility Characteristics segment is constructed using data from the Facility Questionnaire, which provides information about survey-collected facility stays, and the administrative Provider of Service (POS) file, which provides facility characteristics pertaining to SNF stays.

For a beneficiary in the current year's population file, any facility stay within a round from the current file year, as well as from the following winter round, provided that it has an admission date that falls within the current file year, is included in the file. The inclusion of these winter round records is meant to capture any stays which began after the conclusion of the fall round for a given file year. Selected data from the POS file is also included for any SNF stay occurring during the file year for beneficiaries on the finder file.

10.3.14.2 Variable Definitions

Please see the Codebook for information regarding variables in this segment.

10.3.14.3 Special Notes

Special Notes regarding the Integration of CASPER Data with FACCHAR

For beneficiaries for which the facility respondent reported a CCN, 12 variables in FACCHAR are skipped during data collection. The survey-reported data are later merged with CASPER administrative data during data processing using the CCN. The values in the CCN questionnaire lookup tool are derived from CASPER, ensuring that matching administrative data will be available during data processing.

In 2022, 100 percent of FACCHAR records where a CCN was reported were matched to CASPER. A flag variable, D_SOURCE, indicates whether the FACCHAR record has been populated for qualifying variables using CASPER.

For some variables, data from CASPER are not directly comparable to the survey-reported items. FMRBEDS, PCHBED, and HDLICBED cannot be substituted using CASPER during data processing and thus remain missing on FACCHAR.

Services provided by the facility are derived from multiple CASPER variables.

- BATHHELP: If more than one resident needs help from staff for bathing or more than one resident completely depends on staff for bathing, then the facility is classified as providing help with bathing.
- DRESHHELP: If more than one resident needs help from staff for dressing or more than one resident completely depends on staff for dressing, then the facility is classified as providing help with dressing.
- EATHELP: If more than one resident needs help from staff for eating or more than one resident completely depends on staff for eating, then the facility is classified as providing help with eating.
- NORMCARE: The facility is classified as providing nursing or medical care if more than one resident receives any of the following types of care: a catheter, radiation therapy, chemotherapy, dialysis, intravenous therapy, respiratory treatment, tracheostomy care, ostomy care, suctioning, injections, or tube feedings.
- SUPRMEDI: The facility is classified providing supervision over medications if more than one resident receives any of the following types of medications: psychoactive, antipsychotic, antianxiety, antidepressant, hypnotic, antibiotics, or pain management.

The FACCHAR variables that may be populated during data processing using CASPER data are: BATHHELP, CAIDBEDS, CANDCBED, CAREBEDS, and DRESHELP. D_TOTBED, which captured the total bed count, is not available in 2022.

10.3.15 Falls (FALLS)

10.3.15.1 Core Content

The Falls segment contains responses related to injuries and attitudes related to falls.

10.3.15.2 Variable Definitions

Please see the Codebook for information regarding variables in this segment.

10.3.15.3 Special Notes

The "other specify" question FALOTHOS is back coded as necessary into the "type of injury from fall" response options, but the verbatim text is not released.

10.3.16 Food Insecurity (FOODINS)

10.3.16.1 Core Content

The Food Insecurity segment contains information regarding the beneficiary's access to sufficient food. These questions are part of the IAQ and are based upon the USDA ERS Six-Item Short Form of the Food Security Survey Module found at <https://www.ers.usda.gov/topics/food-nutrition-assistance/food-security-in-the-us/survey-tools>.

10.3.16.2 Variable Definitions

Please see the Codebook for information regarding variables in this segment.

10.3.16.3 Special Notes

This questionnaire is administered the summer following the year of interest. The food insecurity section for the reference year 2022 was asked in the summer of 2023. Special non-response adjustment weights are included in the file to account for survey non-response from the fall to the summer data collection period. Note that counts of cases with positive Topical weights may vary within the data year and may change across years due to response rates, sample sizes, and fielding methods. The Topical weights account for these changes. Please see section 9.4.2 for information on using Topical weights.

10.3.17 General Health (GENHLTH)

10.3.17.1 Core Content

The General Health segment contains data regarding a beneficiary's general health status and functioning such as height and weight.

10.3.17.2 Variable Definitions

BMI_CAT: BMI (Body Mass Index) was calculated using height and weight as-

$$(WEIGHT*703)/((HEIGHTFT*12+HEIGHTIN)*(HEIGHTFT*12+HEIGHTIN))$$

Then categorized as:

$0 < \text{BMI} < 18.5 = 1$
 $18.5 \leq \text{BMI} < 25 = 2$
 $25 \leq \text{BMI} < 30 = 3$
 $30 \leq \text{BMI} < 40 = 4$
 $\text{BMI} \geq 40 = 5$

10.3.17.3 Special Notes

For height and weight information at HEIGHTFT, HEIGHTIN, and WEIGHT, the respondent is asked to recall or estimate, not to measure or weigh him or herself. In the height measurement, fractions of an inch have been rounded: those one-half inch or more were rounded up to the next whole inch, those less than one-half inch were rounded down. In the weight measurement, fractions of a pound have been rounded: those one-half pound or more were rounded up to the next whole pound, those less than one-half pound were rounded down.

In 2022, the code frame at VITSUPYR, which collects what vitamins and dietary supplements the beneficiary has taken during the reference period, was updated to include a "Not applicable" response for beneficiaries who only take a multivitamin.

10.3.18 Health Insurance Summary (HISUMRY)

10.3.18.1 Core Content

The Health Insurance Summary segment contains information on administrative plans and their characteristics. Specifically, it includes flags for monthly enrollment and dual eligibility status and information on premiums, co-pays, deductibles, and capitated payments. The file also includes EST_TPRM, which is the sum of premiums for Parts A, B, C, and D and premiums for other plans (private coverage purchased directly from an insurance company, etc.).

There are important caveats to using premium information contained in HISUMRY. For more details, see the notes below on the H_PDLS01-12: Low-Income Subsidy Indicator values.

10.3.18.2 Variable Definitions

H_DUAL01-12: The variables H_DUAL01-H_DUAL12 describe dual eligibility for each month based on state reporting requirements outlined in the Medicare Prescription Drug, Improvement, and Modernization Act of 2003 (MMA). These variables provide more detail regarding the type of Medicaid benefits the beneficiary is entitled to receive and are considered the most accurate source of information on enrollee status. Specific types of dual eligibility identified by these variables are as follows, where the applicable month is MM:

- Qualified Medicare Beneficiaries without other Medicaid (QMB-only) – These individuals are entitled to Medicare Part A, have an income of 100 percent of the Federal poverty level (FPL) or less, have resources that do not exceed twice the limit for SSI eligibility, and are not otherwise eligible for full Medicaid. Medicaid pays their Medicare Part A premiums, if any, Medicare Part B premiums, and to the extent consistent with the Medicaid State plan, Medicare deductibles, and coinsurance for Medicare services provided by Medicare providers. [Partial benefit; H_DUALMM=01]
- Qualified Medicare Beneficiaries plus full Medicaid (QMB-Plus) – These individuals are entitled to Medicare Part A, have an income of 100 percent FPL or less and resources that do not exceed twice the limit for SSI eligibility, and are eligible for full Medicaid benefits. Medicaid pays their Medicare Part A premiums, if any,

Medicare Part B premiums, and, to the extent consistent with the Medicaid State plan, Medicare deductibles and coinsurance, and provides full Medicaid benefits. [Full benefit; H_DUALMM=02]

- Specified Low-Income Medicare Beneficiaries without other Medicaid (SLMB-only) – These individuals are entitled to Medicare Part A, have an income of greater than 100 percent FPL but less than 120 percent FPL, have resources that do not exceed twice the limit for SSI eligibility, and are not otherwise eligible for Medicaid. Medicaid pays their Medicare Part B premiums only. [Partial benefit; H_DUALMM=03]
- Specified Low-Income Medicare Beneficiaries plus full Medicaid (SLMB-Plus) – These individuals are entitled to Medicare Part A, have an income of greater than 100 percent FPL but less than 120 percent FPL, have resources that do not exceed twice the limit for SSI eligibility, and are eligible for full Medicaid benefits. Medicaid pays their Medicare Part B premiums and provides full Medicaid benefits. [Full benefit; H_DUALMM=04]
- Qualified Disabled and Working Individuals (QDWI) – These individuals lost their Medicare Part A benefits due to their return to work. They are eligible to purchase Medicare Part A benefits, have an income of 200 percent FPL or less, have resources that do not exceed twice the limit for SSI eligibility, and are not otherwise eligible for Medicaid. Medicaid pays the Medicare Part A premiums only. [Partial benefit; H_DUALMM=05]
- Qualifying Individuals (QI) – There is an annual cap on the amount of money available, which may limit the number of individuals in the group. These individuals are entitled to Medicare Part A, have an income of at least 120 percent FPL but less than 135 percent FPL, have resources that do not exceed twice the limit for SSI eligibility, and are not otherwise eligible for Medicaid. Medicaid pays their Medicare Part B premiums only. [Partial benefit; H_DUALMM=06]
- Other full benefit dually eligible/Medicaid Only Dually Eligibles (Non-QMB, -SLMB, -QDWI, -QI) – These individuals are entitled to Medicare Part A and/or Part B and are eligible for full Medicaid benefits. They are not eligible for Medicaid as a QMB, SLMB, QDWI, or QI. Medicaid provides full Medicaid benefits and pays for Medicaid services provided by Medicaid providers, but Medicaid will only pay for services also covered by Medicare if the Medicaid payment rate is higher than the amount paid by Medicare, and, within this limit, will only pay to the extent necessary to pay the beneficiary's Medicare cost sharing liability. Payment by Medicaid of Medicare Part B premiums is a state option. [Full benefit; H_DUALMM=08]

H_DOT: Medicare entitlement end date from the Medicare Administrative data. If the date is beyond the calendar year, it is shown as missing.

H_EGWP01-H_EGWP12: Prescription Drug Plan (PDP) Employer Group Waiver Plan (EGWP) Indicator: If the plan is an EGWP, then the value is 1, else the value is 2. An EGWP is not open to general enrollment but is offered through an employer group.

H_ESREND: Ending date of ESRD period. If the date is beyond the calendar year, then it is shown as missing.

H_GHPSW: Some MCBS beneficiaries belong to Medicare managed care plans. CMS derives variables that describe this Medicare managed care membership (H_GHPSW and H_MAFF01-MAFF12). The variable (H_GHPSW) should be used only when there is an indication that the enrollee was a member of a Medicare managed care plan at some time during 2022, and this information is needed for analysis. The monthly variables (H_MAFF01-H_MAFF12) can be used for analyzing membership at specific points in time. The variables will indicate either "FF" (Original Medicare/Fee for Service), "MA" (Medicare Advantage/Other Medicare Capitated Payment Plans), or "NO" (No Entitlement). The H_GHPSW variable is derived from the Health Maintenance Organization (HMO) Coverage Months variable in the administrative data. This variable indicates participation in a group health organization, also known as HMO, managed care participation, or Medicare Advantage/Medicare Part C.

H_MAFF01-12: The MA flag variables are the most reliable indicators for monthly MA information. This information is sourced from the CMS administrative data. The H_ENT variables were used to determine if the individual did not have Medicare entitlement. This information factored into the "No Entitlement" category in the MA flag monthly variables. The monthly entitlement variables can be found on the HITLINE segment. H_DOE and H_DOT on the HISUMRY file provide Medicare entitlement start and end dates for the beneficiary. Because the administrative source of this information has changed, H_ENT variables cannot be used to crosswalk to the MA flag variables. However, H_ENT can be used to determine Part A and Part B eligibility among FFS beneficiaries in files prior to 2015.

H_MAPMT: Total MA A/B Payment – annual amount from Medicare Advantage Prescription Drug (MARx) data.

H_MCSW: State buy-in is tracked by CMS and used as a general proxy for Medicaid participation. CMS derived H_MCSW using its administrative enrollment data.

H_OPMDCD: This variable provides a summary of annual Medicare-Medicaid dual eligibility based on the state Medicare Modernization Act (MMA) files.

Beneficiaries are assigned a dually eligible status if they are Medicaid eligible for at least one month. Specific eligibility (full, partial, or QMB) is determined by the beneficiary's status in the last month of eligibility for the year (for definitions, see option C below in Special Notes for HISUMRY for Full-benefit vs. Partial-benefit vs. QMB-only). QMB beneficiaries include Qualified Medicare Beneficiaries without other Medicaid (QMB-only). The "partial benefit" beneficiaries include: Specified Low-Income Medicare Beneficiaries without other Medicaid (SLMB-only), Qualified Disabled and Working Individuals (QDWI), and Qualifying Individuals (QI). The "full benefit" beneficiaries include Qualified Medicare Beneficiaries plus full Medicaid (QMB-Plus), Specified Low-Income Medicare Beneficiaries (SLMB-Plus), and all other full benefit beneficiaries (Non-QMB, -SLMB, -QWDI, -QI).

Medicaid Questions: To help the respondent answer the questions about Medicaid, the interviewers use the name of the Medicaid program in the state where the beneficiary lives. A health insurance plan is one that covers any part of hospital bills, doctor bills, or surgeon bills, but does not include any of the following:

- Public plans, including Medicare and Medicaid, mentioned elsewhere in the questionnaire.
- Disability insurance which pays only on the basis of the number of days missed from work.
- Veterans' benefits.
- "Income maintenance" insurance which pays a fixed amount of money to persons both in and out of the hospital or "Extra Cash" policies. These plans pay a specified amount of cash for each day or week that a person is hospitalized, and the cash payment is not related in any way to the person's hospital or medical bills.
- Workers' Compensation.
- Any insurance plans that are specifically for contact lenses or glasses only. Any insurance plans or maintenance plans for hearing aids only.
- Army Health Plan and plans with similar names (e.g., CHAMPUS, CHAMPVA, Air Force Health Plan).
- Dread disease plans that are limited to certain illnesses or diseases such as cancer, stroke, or heart attacks.
- Policies that cover students only during the hours they are in school, such as accident plans offered in elementary or secondary schools.
- Care received through research programs such as the National Institutes of Health.

H_PDLS01-12: Low-Income Subsidy Indicator values: When conducting data analysis with the variables H_PDLS01-12 from the 2013 and earlier files and the 2015 and later files, users will need to recode the 2015 and later data to the previous values. See the *2018 and 2019 Data User's Guides: Survey File* for recoding guidance. The eligible categories are beneficiaries who are deemed eligible, and these beneficiaries are automatically enrolled.

H_DDED01-12: The monthly values reflect the annual Part D deductible amount charged by the plan that the beneficiary was enrolled in that month.

H_PRPY01-04: Primary Payer codes are summarized from the FFS claims. These codes indicate that some other payer besides Medicare covered at least some portion of the charges. Additional detail can be found under NCH_PRMRY_PYR_CD, <https://www2.ccwdata.org/documents/10280/19022436/codebook-ffs-claims.pdf>

H_PTAPRM: Total Part A Premium paid in calendar year (CY) – This is for beneficiaries who purchased Part A by paying a monthly premium. Note that this variable will have a relatively small number of beneficiaries.

H_PTBPBM: Total Part B Premium paid in CY – This includes all Part B beneficiaries (a large number; a premium is always paid by either the beneficiary or a third party). NOTE: The MCBS shows no Part B premium paid if the beneficiary belongs to a managed care plan in which the plan pays the entirety of the premium. In this scenario, the plan paid the entirety of the beneficiary's premium, so the process shows no premium paid.

H_PTDAMT: PTD Total Payment – annual amount from the MARx data.

EST_TPRM: This variable is the sum of all premiums reported prorated by the number of months of coverage for each plan. The variable name emphasizes that the total is an estimate since complete information on the amount that a beneficiary paid may not be available for all plans. For Part A, B, C, and D plans, the premium reflects the total paid, either by the beneficiary or a third party on their behalf. Starting with 2022, premium information for private plans was imputed for beneficiaries who did not provide a response. This may affect the total value of EST_TPRM for some records. The Part C portion is based on premiums reported to CMS, and it is important to note that this information may not be available, as in the case of plans obtained through a beneficiary's employer. Additionally for Part C, the EST_TPRM is based on the administrative premiums only (administrative and survey-reported data are not reconciled). Due to the availability of premium data, it is important to note that EST_TPRM is an estimate.

Payment Model Participation Flags

There are three variables that indicate the payment model for each plan.

H_PRGID: CMS Prog ID – Payment Model

H_PRGID2: 2nd CMS Prog ID – Payment Model

H_PRGID3: 3rd CMS Prog ID – Payment Model

H_PRGID2 and H_PRGID3 are only populated if the beneficiary has multiple program IDs. Variables are designated as single, 2nd, or 3rd based on the start/end dates of the entries in the source data (earliest start date, next=2, etc.). Start dates are prior to 12/31/YR and end dates may be after 1/1/YR where "YR" = data year.

10.3.18.3 Special Notes

When describing dually eligible enrollees, users typically define and present analyses separately for two subgroups: full-benefit and partial-benefit. However, some users may wish to pull the QMB-only beneficiaries

out of the partial-benefit group to create a third classification. Therefore, the H_DUAL01-H_DUAL12 variables may be used to group Medicare-Medicaid enrollees into one, two or three categories, as follows:

A. No delineation:

All Medicare-Medicaid (dually eligible) enrollees: H_DUAL01-H_DUAL12 in (01, 02, 03, 04, 05, 06, 08)

B. Full-benefit vs. Partial-benefit:

Partial-benefit: H_DUAL01-H_DUAL12 in (01, 03, 05, 06)

Full-benefit: H_DUAL01-H_DUAL12 in (02, 04, 08)

C. Full-benefit vs. Partial-benefit vs. QMB-only:

QMB-only: H_DUAL01-H_DUAL12 =01

Partial-benefit (non-QMB): H_DUAL01-H_DUAL12 in (03, 05, 06)

Full-benefit: H_DUAL01-H_DUAL12 in (02, 04, 08)

For detailed information on how the HITLINE and HISUMRY segments differ from the previously released RICs (i.e., RICs 4 and A), see the *2018 and 2019 Data User's Guides: Survey File*.

10.3.19 Health Insurance Timeline (HITLINE)

10.3.19.1 Core Content

The Health Insurance Timeline segment contains one record for each plan a beneficiary has and includes information on type of insurance coverage, monthly eligibility/enrollment, coverage start and end dates, and the source information for the coverage. For all plans that a beneficiary has, both administrative and survey reported are included on the file. However, starting with 2021, survey reports of Medicare Advantage (MA) enrollment with no corresponding record of MA enrollment in administrative data have been excluded. In addition, HITLINE contains detailed information on plans for which no administrative data are available. These plans are reported in the survey only and include different types of private plans, Tricare, coverage for certain medical events through the Department of Veteran's Affairs for beneficiaries living in a facility, and public plans that do not fall under either Medicare or Medicaid. For these survey-only plans, the file includes flags indicating types of services covered, and, for private plans, information on plan policyholder and premiums paid. All plans reported in a Community setting also have a unique plan identifier, PLANNUM, which can be used to link plans across multiple years.

The questionnaire does not ask whether a given plan offers 'comprehensive' coverage. Data users can construct their own definition of comprehensive coverage and consult individual coverage flags to determine if a plan meets their criteria for being a comprehensive plan.

10.3.19.2 Variable Definitions

SRCCOV01-12: Indicates the source of coverage information for the plan for a given month in the calendar year: CMS Administrative Data, Survey Data, or Both Administrative and Survey Data.

COV01-12: Indicates if the beneficiary was covered by this plan for a given month in the calendar year.

S_DVH: Indicates whether plan covers dental, vision, or hearing services.

S_OTHPLN: Indicates whether plan is a specialty plan that only covers specific services (such as long-term care, coverage for cancer/dread disease, etc.).

S_HMOPPO: Indicates whether beneficiary's private plan is an HMO/PPO. Obtained from the HIQ variable PPRVHMO.

S_PHREL: The relationship of the policyholder to the beneficiary. Responses from the HIQ variable PERS_MIPNUM are combined with beneficiary's household roster information to determine the policyholder's relationship to the beneficiary.

S_OBTNP: Indicates how the main insured person obtained their private policy (e.g., self-purchased, current or former employer, etc.). Obtained from the HIQ variable PPRVGET.

S_COVNM: The number of people covered by each private plan. Obtained from the HIQ variable PRVNMCOV.

D_COVRX: Indicates if beneficiary's plan covers prescription drugs.

S_MSCOV: Indicates if beneficiary's plan covers visits to a doctor or other professional or lab work. Obtained from the HIQ variable PRVMSCOV.

S_IP: Indicates whether beneficiary's private plan covers inpatient stays. Obtained from the HIQ variable PRVIPCOV.

S_COVNH: Indicates whether beneficiary's private plan has long-term care coverage. Obtained from the HIQ variable PRVNHCOV.

S_DNTAL: Indicates whether beneficiary's private plan covers dental services. Obtained from the HIQ variable MHMODENT.

S_VISN: Indicates whether beneficiary's private plan covers optical or vision coverage. First added in Fall 2020 (Round 88), this information is obtained from the HIQ variable PRVOPEYE.

S_PAYSP: Indicates whether the main insured person (MIP) pays any part of the insurance premium. Obtained from the HIQ variable MIPPINS.

D_PREMMON: Monthly cost of private health insurance plan premiums. A premium amount was recorded even if the respondent did not directly pay the premium (for example, if a son or daughter paid the premium). This variable was derived from the HIQ variable MIPPAMT. For family plans, the reported amount reflects the total premium paid for the plan. Reported premiums were converted to monthly values. For beneficiaries who did not provide complete premium information, premium information was imputed.

D_PREMN_I: Imputation flag for D_PREMMON.

S_PAYOTH: Indicates whether anyone else, such as an employer or a union, helped to pay any portion of the premium. Obtained from the HIQ variable MHMOCOST.

S_PAYWHO: Indicates who paid a portion of the total cost of the premium. Obtained from the HIQ variable MHMOWHO.

S_TRIRX: Specifies where Tricare members obtain prescription drugs. Obtained from the HIQ variable TRIMEDS.

D_FCLTYF: Indicates whether a plan was reported in a Facility setting. Facility interviews are not conducted with the beneficiary but rather with facility staff who may have little information on coverage type and plan details. D_FCLTYF indicates which plans were reported in a Facility setting and thus have limited detailed information about them available. Beneficiaries who transition between community and facility settings may

have a plan reported in each setting. However, due to the nature of the Facility interview, it is not possible to ascertain whether these would reference the same plan.

10.3.19.3 Special Notes

The HITLINE segment has one record for every plan reported for a beneficiary. Individuals covered for the entire year by a plan will have a BEGDATE of 010120XX and an ENDDATE of 123120XX to indicate a full year's coverage. BEGDATE is set for all plans using the month when a plan was first reported. For example, if someone had coverage January – March and June – November, BEGDATE will reflect that coverage started in January. Most plans have an ENDDATE as well. The only plans with missing ENDDATE are plans where coverage ended and then started again. For plans where survey and administrative data are combined, BEGDATE and ENDDATE are set using all available coverage information. Data users can reference SRCCOV01-SRCCOV12 flags to identify whether coverage information for a given month came from administrative records, a survey report, or both.

Eligibility for Tricare can be lost. Due to this fact, data users should pay attention to the appropriate coverage indicators (i.e., PLANTYPE, COV01-COV12).

Beginning in 2022, D_PREMMON was added to capture monthly insurance plan premiums for private plans and is imputed for private plans with missing premium amount values. An imputation flag variable (D_PREMN_I) was also added to identify the imputed values on D_PREMMON and the previous premium variables on HITLINE were dropped (S_PAYUNT, S_PREM, and D_ANNPRM). Information about MA plan premiums is available on MAPLANQX.

In Winter 2023, several updates were made to redesign the HIQ section. As a result, health insurance plans reported in that round under the new questionnaire structure with coverage extending back to 2022 are included in the 2022 Survey File. The redesigned structure also obtains more accurate information on the start and end dates of health insurance plan coverage, which makes it easier to determine whether a plan should be included in the 2022 Survey File or the 2023 Survey File only.

10.3.20 Household Characteristics (HHCHAR)

10.3.20.1 Core Content

The Household Characteristics segment includes beneficiaries who resided in a community setting as of their last complete interview and contains information about the beneficiary's household composition and residence. For each calendar year, this segment reflects the latest available data on the size of the household and the age and relationship of household members. Information about the beneficiary's physical residence is collected at the Baseline interview and updated as necessary.

10.3.20.2 Variable Definitions

D_HHTOT: Reflects the total number of people living in the household.

D_COMPHH: Reflects the composition of the household members.

D_SEXSPP: Indicates the sex when a spouse or partner is identified in D_COMPHH as a member of the household.

D_HHLT50 and D_HHGE50: Indicate the number of people in the household under the age of 50 and people 50 years of age or above, respectively. These numbers may include the beneficiary.

D_HHLT18: Indicates the number of people under the age of 18 who are related to the beneficiary.

10.3.20.3 Special Notes

CMS defines a household as a group of individuals, either related or not, who live together and share one kitchen. This may be one person living alone, a head of household and relatives only, or a head of household living with relatives, boarders, and any other unrelated individual living under the same roof, sharing the same kitchen.

Household membership includes all persons who currently live at the household or who normally live there but are away temporarily. For example, unmarried students away at school or family members away receiving medical care are included. Visitors in the household who will be returning to a different home at the end of the visit are not included. Generally, if there is any question about the composition of the household, the respondent's response is accepted.

Because the date of birth is sometimes unknown (perhaps because a proxy provided the information), the sum of the variables "number under 50"/"number 50 or older" (D_HHLT50/D_HHGE50) may not equal the total number of people in the household (D_HHTOT).

Data on certain characteristics of the residence (e.g., number of levels) is collected during the Baseline interview and carried forward unless a beneficiary moved or had a Facility stay prior to returning to the Community. Information about other characteristics of the residence (e.g., availability of personal care services) is updated annually during the fall interview.

Only beneficiaries living in the community who are responding to a Continuing interview are in universe for the question SPMOVED, "Has the SP moved since the last Fall Round data collection date?". For this reason, data users are encouraged to use longitudinal weights if they wish to utilize this variable in analysis. The reference period for this variable is going to be longer for beneficiaries whose last fall interview was in a facility and beneficiaries who missed the last fall interview.

The "other specify" questions DWELLOS and HCOMUNOS are back coded as necessary into the "description of beneficiary's housing" response options, but the verbatim text is not released.

The questionnaire was also modified to only collect full date of birth for the beneficiary and their spouse or partner. Therefore, logic was added at variable EHHDOBDD to hide the "day" field if the household member is anyone other than the beneficiary's partner or spouse.

In 2022, the variables D_HHREL and D_HHUNRL were dropped from HHCHAR as they could not accurately count the number of related and unrelated individuals in the household.

10.3.21 Income and Assets (INCASSET)

10.3.21.1 Core Content

This segment contains data on a beneficiary's reported income and assets.

10.3.21.2 Variable Definitions

Please see the Codebook for information regarding variables in this segment.

10.3.21.3 Special Notes

In the IAQ, for a few income questions the reference period is the previous calendar year. That is, these income questions are asked in the summer of 2023 about income earned in 2022.

Example: "Now I want to ask about your [and spouse's or partner's] total income for last year, that is, for the calendar year ending in December [CURRENT YEAR - 1], before any federal or state taxes were taken out."

Other items ask about income earned in the current calendar year.

Example: "You told me earlier that you have job-related pension plans. In all, how much was received from these pension plans in the last month, before any federal or state taxes were taken out (for the month of [CURRENT MONTH - 1])?"

For assets, there are three different timeframes referenced in the IAQ:

1. How much of an asset was received or withdrawn in the last month.
 - a. Example: "Is your mortgage paid off or are monthly mortgage payments still being made?"
2. How much is currently in certain accounts.
 - a. Example: "This next question is a bit different. You mentioned that you have retirement accounts. In total, about how much is currently in all of these retirement accounts?"
3. How much altogether was received or withdrawn in the last year.
 - a. Example: "Now thinking about all of last year, that is calendar year [CURRENT YEAR - 1], how much altogether did you receive or withdraw from all of these retirement accounts?"

The difference in reference periods between income and assets items is due to the nature of the information collected (i.e., respondent recall is facilitated when asking about a bank account balance from the last month versus four months ago), and many assets are relatively stable in value (e.g., housing).

Special non-response adjustment weights are included in the file to account for survey non-response from the fall to the summer data collection period. Note that counts of cases with positive Topical weights may vary within the data year and may change across years due to response rates, sample sizes, and fielding methods. The Topical weights account for these changes. Please see section 9.4.2 for information on using Topical weights.

The MCBS imputes income when income data are missing. Data are first imputed whether or not an income source (such as Social Security) exists. If the income source exists, then the amount earned is imputed next. Imputation is performed using the hot deck imputation method, and a flag is created for each imputed variable indicating whether or not the corresponding value is imputed.

The "other specify" questions LUMPSUMO and OPYSCHED are back coded as necessary into the "form of lump payment" and "other payment schedule" response options, respectively, but the verbatim text is not released.

In 2022, the INCASSET imputation added separate variables for monthly earnings from work for the beneficiary and for their spouse/partner. Previously, these variables were combined into a single variable and this variable was imputed on its own.

10.3.22 Interview Characteristics (INTERV)

10.3.22.1 Core Content

The Interview Characteristics segment summarizes interview characteristics, such as the type of interview and whether a proxy is used.

10.3.22.2 Variable Definitions

Please see the Codebook for information regarding variables in this segment.

10.3.22.3 Special Notes

Some beneficiaries have more than one interview in a round if they have moved from a community to facility setting within the reference period, or vice versa. To avoid duplication of data, the information in this file represents the last interview conducted with the respondent in each given round. INTERVU indicates which type of interview (Community or Facility) was conducted.

TOTLINTV indicates the total number of interviews conducted with the beneficiary in the year.

In 2022, the code frame for INTMODE was updated to include hybrid interviews conducted using a mix of phone and in-person interviewing. See section 10.1.8 for more information on the interview mode indicator.

10.3.23 Medicare Advantage Plan Questions (MAPLANQX)

10.3.23.1 Core Content

The MA Plan Questions segment augments information from the ACQ and SCQ sections of the questionnaire for beneficiaries enrolled in Medicare Part C. Beneficiaries who are enrolled in an MA plan at the time of the interview are asked general questions about their health plans, which include access to and satisfaction with medical services. This segment also contains the beneficiary's assessment of the quality of the medical care that they are receiving, types of additional coverage offered, and any beneficiary-paid premiums associated with the health plan.

10.3.23.2 Variable Definitions

D_ANHMO: The annual additional cost of MA premiums. The premiums are annualized regardless of the length of time the respondent actively held the policy. This variable is derived from the HIQ items MHMOAMT and MHMOUNIT.

ANHMO_I: The imputation flag for D_ANHMO.

D_MADV: This variable is derived from administrative data and set to 1 if the beneficiary was covered by an MA plan for at least one month out of the calendar year. All beneficiaries included in the MA Plan Questions segment have D_MADV set to 1.

MADVYRS: The number of years the beneficiary has been enrolled in MA. This variable is derived from the HIQ item HMONUMYR.

RECMADV: Indicates whether the respondent recommends the MA plan to family/friends. This variable is derived from the HIQ item RECMHMO.

MA coverage variables (MADVNT, MADVNH, MADVEYE, and MADVRX): Indicate whether the beneficiary's MA plan covers dental care, vision care, nursing home care, and prescription medicines. These variables are derived from the HIQ items MHMODENT, MHMONH, MHMOEYE, and MHMORX.

MA payment variables (MADVPAY, MADVCOST, and MADVWHO): Indicate whether there is an additional cost associated with the MA plan and if so, who covers the cost. These variables are derived from the HIQ items MHMOPAY, MHMOCOST, and MHMOWHO.

10.3.23.3 Special Notes

If the respondent reports a payer or a unit of payment that is not included in the predefined code list, the interviewer documents their response verbatim in an "other specify" variable that is not released. The "other specify" response is back coded as necessary into the predefined code list.

Beginning in 2022, D_ANHMO (annual additional cost of MA premiums) includes imputed premiums for beneficiaries who did not provide complete premium information about their MA plan. As such, an imputation flag variable (ANHMO_I) was added to identify the imputed values on D_ANHMO. Information about private plan premiums is available on HITLINE.

10.3.24 Medicare Plan Beneficiary Knowledge (MCREPLNQ)

10.3.24.1 Core Content

The Medicare Plan Beneficiary Knowledge segment contains information from the KNQ section related to the beneficiary's knowledge about the Medicare open enrollment period and Medicare-covered expenses. The KNQ is administered the winter following the year of interest.

The data collected in this segment support evaluation of the impact of existing education initiatives by CMS. The KNQ section helps refine future CMS education initiatives by asking about information that beneficiaries may need, preferred sources for this information, and beneficiaries' access to insurance information. This data also presents the knowledge beneficiaries have gained from CMS publications.

10.3.24.2 Variable Definitions

KVSTSITE: This variable collects whether the respondent has ever visited the official website for Medicare information. If the respondent has previously answered "yes" to this question, the "yes" response is pulled forward to the current data year.

KCPHINFO: This variable collects whether the respondent has ever called 1-800-MEDICARE. If the respondent has previously answered "yes" to this question, the "yes" response is pulled forward to the current data year.

10.3.24.3 Special Notes

In Winter 2023, four new variables were added to the KNQ section. Two items, USEMSP and APPLYMSP, were added to gather beneficiary participation in the Medicare Savings Program (MSP). RGHTAPL was added to understand beneficiary knowledge of the Medicare appeal process.

INTERNET, which quantifies beneficiary access to the internet, was also added to the KNQ section in Winter 2023. As a result, the universes of the variables KNETPERS (beneficiary internet usage to obtain information), KNETFRND (if the beneficiary has someone who gets information from the internet for them), and KNETOFTN (how often the beneficiary accesses the internet to get information) were modified to only be administered if

the beneficiary reports access to the internet at INTERNET. For that reason, the variable names of KNETPERS, KNETFRND, and KNETOFTN were changed to USENET, SOMELNET, and OFTNNET, respectively.

Special non-response adjustment weights are included in the file to account for survey non-response from the fall to the winter data collection period. Note that counts of cases with positive Topical weights may vary within the data year and may change across years due to response rates, sample sizes, and fielding methods. The Topical weights account for these changes. Please see section 9.4.2 for information on using Topical weights.

10.3.25 Minimum Data Set (MDS3)

10.3.25.1 Core Content

The Minimum Data Set is health assessment information collected while the beneficiary was in an approved Medicare Facility. For more information regarding the MDS and the changes in version 3.0, please consult <https://www.cms.gov/medicare/quality/nursing-home-improvement>.

10.3.25.2 Variable Definitions

Please see the Codebook for information regarding variables in this segment.

10.3.25.3 Special Notes

MDS3 administrative data records are included for any beneficiary having such a record in the year of interest. There are beneficiaries living in the community (DEMO segment INT_TYPE = C) that appear in the MDS segment. CMS includes MDS data for all MCBS beneficiaries regardless of the INT_TYPE, which is determined by the type of survey instrument completed.

For more information on the difference between the MDS and FACASMNT data and how the two segments can be linked, please see the FACASMNT section above.

10.3.26 Mental Health (MENTHLTH)

10.3.26.1 Core Content

The Mental Health segment contains survey responses regarding the beneficiary's mental health such as feelings of anxiety or depression.

10.3.26.2 Variable Definitions

Please see the Codebook for information regarding variables in this segment.

10.3.26.3 Special Notes

Generalized Anxiety Disorder scale (GAD-2): Two items labeled with "GAD" comprise the GAD-2 scale, which is a screening tool for generalized anxiety.

Patient Health Questionnaire (PHQ-9): Items labeled with "PHQ" are taken from the PHQ-9, which is a screening tool for depression. The MCBS does not collect the ninth item on the PHQ-9, which asks about suicidal ideation, but does include the PHQ-9 follow-up question that asks about the overall difficulty caused by depression (MENTHLTH item PHQPRDIF).

10.3.27 Mobility (MOBILITY)

10.3.27.1 Core Content

The Mobility segment contains information on the beneficiary's use of available transportation options and whether the beneficiary's health affects their daily travel.

10.3.27.2 Variable Definitions

Please see the Codebook for information regarding variables in this segment.

10.3.27.3 Special Notes

N/A

10.3.28 Multiple Year Enrollment (MYENROLL)

10.3.28.1 Core Content

The Multiple Year Enrollment segment combines five years of enrollment information for the current year MCBS beneficiary population. This allows users to view multiple years of enrollment information in one file.

10.3.28.2 Variable Definitions

ENROLYR: Indicates the enrollment data year.

H_MAFF01-12: These MA flag variables are the most reliable indicators for monthly MA information. This information is sourced from CMS administrative data.

PTA_MONS: Indicates the number of months the beneficiary had Medicare Part A Hospital Insurance (HI) coverage.

PTB_MONS: Indicates the number of months the beneficiary had Medicare Part B Supplemental Medical Insurance (SMI) coverage.

H_PTD01-12: These Part D plan flags indicate whether there was Medicare Part D coverage for the month.

H_DUAL01-12: These variables indicate the dual eligibility code by month.

H_PDLS01-12: These variables indicate the Low Income Subsidy (LIS) Indicator by month.

10.3.28.3 Special Notes

N/A

10.3.29 Nagi Disability (NAGIDIS)

10.3.29.1 Core Content

The Nagi Disability segment contains information on the beneficiary's difficulties with performing ADLs and IADLs, including which ADLs and IADLs the beneficiary has difficulty performing, how long the beneficiary has experienced these difficulties, whether the beneficiary has received any help or used supportive equipment to perform ADLs or IADLs, and the total number of persons who have helped the beneficiary, if applicable.

10.3.29.2 Variable Definitions

ADL and IADL measures: The MCBS asks respondents whether they have any difficulty performing 12 activities. Their answers about difficulty performing the IADLs (PRBTELE, PRBLHWK, PRBHHWK, PRBMEAL, PRBSHOP, and PRBBILS) and ADLs (HPPDBATH, HPPDDRES, HPPDEAT, HPPDCHAR, HPPDWALK, and HPPDTOIL) reflect whether or not the beneficiary usually has difficulty and anticipates continued trouble with these tasks, even if a short-term injury made them temporarily difficult.

"Difficulty" in these questions has a qualified meaning. Only difficulties associated with a health or physical problem are considered. If a beneficiary only performed an activity with help from another person (including just needing to have the other person present while performing the activity), then that respondent is deemed to have difficulty with the activity.

Help from another person includes a range of helping behaviors. The concept encompasses personal assistance in physically doing the activity, instruction, supervision, and "standby" help. These questions are asked in the present tense; the difficulty may have been temporary or may be chronic. Vague or ambiguous answers, such as "Sometimes I have difficulty," are coded "yes."

D_ADLHNM: D_ADLHNM stores the number of persons helping the beneficiary with ADLs and/or IADLs. D_ADLHNM is derived by counting the number of helper rows for a BASEID.

D_MODTIM: The length of time the beneficiary spent doing moderate activities (e.g., golf, gardening) is collected in number of minutes/day, hours/day, hours/week, or hours/month. The length of time is stored in a continuous variable while the corresponding unit is stored in a categorical variable. These variables are used to derive D_MODTIM, the number of hours per week the beneficiary spent doing moderate activities.

D_MUSTIM: The length of time the beneficiary spent increasing muscle strength (e.g., lifting weights, yoga) is collected in number of minutes/day, hours/day, hours/week, or hours/month. The length of time is stored in a continuous variable while the corresponding unit is stored in a categorical variable. These variables are used to derive D_MUSTIM, the number of hours per week the beneficiary spent increasing muscle strength.

D_VIGTIM: The length of time the beneficiary spent doing vigorous activities (e.g., running, aerobics) is collected in number of minutes/day, hours/day, hours/week, or hours/month. The length of time is stored in a continuous variable while the corresponding unit is stored in a categorical variable. These variables are used to derive D_VIGTIM, the number of hours per week the beneficiary spent doing vigorous activities.

HPPDBATH: Beneficiaries who have difficulty bathing or showering without help met at least one of the following criteria:

- someone else washes at least one part of the body
- someone else helps the person get in or out of the tub or shower or helps get water for a sponge bath
- someone else gives verbal instruction, supervision, or stand-by help
- the person uses special equipment such as handrails or a seat in the shower stall
- the person never bathes at all (a highly unlikely possibility)
- the person receives no help, uses no special equipment or aids, but acknowledges having difficulty

HPPDDRES: Dressing is the overall complex behavior of getting clothes from closets and drawers and then putting the clothes on. Tying shoelaces and putting on socks or hose are not considered part of dressing. Special dressing equipment includes items such as button hooks, zipper pulls, long-handled shoe horns, tools

for reaching, and any clothing made especially for accommodating a person's limitations in dressing, such as Velcro fasteners or snaps.

HPPDEAT: A person eats without help if he or she can get food from the plate into the mouth. A person who does not ingest food by mouth (that is, is fed by tube or intravenously) is not considered to eat at all. Special eating equipment includes such items as a special spoon that guides food into the mouth, a forked knife, a plate guard, or a hand splint.

PRBBILS: Managing money refers to the overall complex process of paying bills, handling simple cash transactions, and generally keeping track of money coming in and money going out. It does not include managing investments, preparing tax forms, or handling other financial activities for which members of the general population often seek professional advice.

PRBLHWK and PRBHHWK: The distinction between light housework (PRBLHWK) and heavy housework (PRBHHWK) is made clear by examples. Washing dishes, straightening up and light cleaning represent light housework; scrubbing floors and washing windows represent heavy housework. The interviewer is not permitted to interpret the answer in light of the degree of cleanliness of the dwelling.

PRBMEAL: "Preparing meals" includes the overall complex behavior of cutting up, mixing, and cooking food. The amount of food prepared is not relevant, so long as it would be sufficient to sustain a person over time. Reheating food prepared by someone else does not qualify as "preparing meals."

PRBSHOP: Shopping for personal items means going to the store, selecting the items, and getting them home. Having someone accompany the beneficiary would qualify as help from another person.

PRBTELE: Using the telephone includes the overall complex behavior of obtaining a phone number, dialing the number, talking and listening, and answering the telephone.

10.3.29.3 Special Notes

Six global disability questions are released to comply with HHS guidance DISDECSN, DISWALK, DISBATH, and DISERRND on the NAGIDIS segment. Variables DISHEAR and DISSEE are included on the VISHEAR segment.

For beneficiaries with identified helpers, information about the persons responsible for assisting with the beneficiary's performance of ADLs and IADLs is found in the ASSIST segment.

10.3.30 Nicotine and Alcohol (NICOALCO)

10.3.30.1 Core Content

The Nicotine and Alcohol segment contains information on the prevalence and frequency of alcohol and nicotine use (including cigarettes, e-cigarettes, cigars, pipe tobacco, and smokeless tobacco).

10.3.30.2 Variable Definitions

Please see the Codebook for information regarding variables in this segment.

10.3.30.3 Special Notes

Affirmative responses indicating former or current use of inhaled tobacco products (cigar, cigarette, smokeless tobacco, pipe tobacco, and e-cigarettes) or alcohol are pulled forward to the current data year variables (i.e., CIGARONE, CIGAR50, CIG100, SMKLSONE, PIPEONE, ECIGONE, and ALCLIFE).

10.3.31 Outcome and Assessment Information (OASIS)

10.3.31.1 Core Content

The Outcome and Assessment Information segment contains assessment information conducted while the beneficiary was receiving home health services.

For more information regarding OASIS, please consult <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HomeHealthQualityInits>.

10.3.31.2 Variable Definitions

Please see the Codebook for information regarding variables in this segment.

10.3.31.3 Special Notes

All home health records are included for MCBS respondents for the year of interest.

10.3.32 Patient Activation (PNTACT)

10.3.32.1 Core Content

The Patient Activation segment contains data that can be used to assess the degree to which beneficiaries actively participate in their own health care and the decisions concerning their health care, measuring if beneficiaries receive information about their health and Medicare and if they understand the information in a way that makes it useful.

10.3.32.2 Variable Definitions

Please see the Codebook for information regarding variables in this segment.

10.3.32.3 Special Notes

Special non-response adjustment weights are included in the file to account for survey non-response as these items are only asked of non-proxy respondents.

10.3.33 Preventive Care (PREVCARE)

10.3.33.1 Core Content

The Preventive Care segment provides data on the beneficiary's use of preventive services, including getting a mammogram, Pap smear, prostate screening, diabetes screening, colon cancer screening, blood pressure screening, flu and pneumonia shots, shingles vaccine, and HIV testing.

10.3.33.2 Variable Definitions

Please see the Codebook for information regarding variables in this segment.

10.3.33.3 Special Notes

Select items are collected only in the summer (whether the beneficiary received a pneumonia shot or the shingles vaccine) while the seasonal flu vaccine items are asked in the winter and summer rounds. Several

items are collected only in the fall, including questions about getting a mammogram, Pap smear, prostate test, blood pressure screening, hysterectomy, and HIV testing.

How do I find out what proportion of Medicare beneficiaries received a flu shot in a given calendar year? Flu shot data are available for both Community and Facility components, but data collection and processing methods are different, and the variables are located on different segments in the Survey File LDS. To estimate prevalence of flu shots in a given flu season, data users need the prior data year (e.g., 2021) Survey File for beneficiaries living in the community and the current data year (e.g., 2022) Survey File for beneficiaries living in a facility. Note that the response categories of the FLUSHOT variables are similar across the two components (Yes/No), but the coding values associated with the Yes/No categories in the LDS files are different. For community, Yes = 1 and No = 2, but for facility Yes = 1 and No = 0. In addition, the reference periods differ between the Community and Facility components. Therefore, users need both the prior and current data year Survey File LDS's to estimate the flu shot prevalence for all Medicare beneficiaries for a given flu season.

Exhibit 10.3.33.1: Segment, Questionnaire, and Variable Information for Analyses of 2021-2022 Flu Shot Data

Component	Variable Location	Variable Label	Data Collection Timing
Community	FLUSHOT ON PREVCARE	FLU SHOT FOR LAST WINTER	PVQ in Winter and Summer 2022 and included in 2021 Survey File
Facility	FLUSHOT on FACASMNT	SP HAD A FLU SHOT IN THE PAST YEAR?	HS in Fall 2022 and included in 2022 Survey File

The "other specify" questions MAMNOTHS, PAPNOTHS, PRONOTHS, FLUOTHOS, and FLUSITOS are back coded as necessary into the reason(s) for not getting a mammogram, Pap smear, prostate test, or flu shot or where they got their flu shot, respectively, but the verbatim text is not released.

10.3.34 Residence Timeline (RESTMLN)

10.3.34.1 Core Content

The Residence Timeline segment provides a timeline of each MCBS setting type in which a beneficiary resides over the portion of the year in which they are enrolled in Medicare, as well as any periods associated with FFS inpatient, SNF, or hospice events.

10.3.34.2 Variable Definitions

D_BEG1: Represents the beneficiary's first date of Medicare eligibility within the file year.

D_CODE1: Either identifies a residential setting or for a small number of cases, contains the code "N". The latter only occurs for some Facility respondents who are new to the MCBS survey but were enrolled in Medicare prior to the start of the year. The first interview that these beneficiaries receive only covers back to the date of admission into the facility in which they currently reside. If they were admitted into their current facility after the 1st of the year, it will result in the setting code on their first situation (D_CODE1) having a value of "N".

10.3.34.3 Special Notes

Residential situations are overwritten by all claim events which overlap them, with two exceptions. Hospice events do not overwrite residential situations as this type of utilization is less indicative of a change in setting as it is a change in the level of care being received. These events should instead be considered as occurring concurrently with the beneficiary's identified residential situation. Also, a beneficiary's initial residential status is not overwritten, even when overlapped completely by a claim of any type, in order to provide context as to their original living situation at the start of their timeline.

The total number of setting changes is equal to the sum of MCBS residential status changes (D_NUMRES) and the number of the events corresponding to the above mentioned claim types (D_NUMEVT). Each transition is identified with a code representing the type of setting along with begin and end dates.

The number of variables in the series D_CODEn, D_BEGn, and D_ENDn will correspond to the maximum number of settings in a given year (calculated as D_NUMSIT + D_NUMEVT). At a minimum, each beneficiary has information pertaining to their setting at the beginning of their eligibility period within the year. Residential status situations do not have end dates populated to illustrate that these extend through any claim events which follow until a change in residential status occurs.

10.3.35 RX Medications (RXMED)

10.3.35.1 Core Content

The RX Medications segment augments information from the ACQ and SCQ sections of the questionnaire with information specific to prescription drug coverage collected in the RXQ section. The RXQ covers topics related to knowledge about and experience with Medicare Part D enrollment, options considered when choosing prescription drug coverage, access to prescription drugs, and satisfaction with current prescription drug coverage.

10.3.35.2 Variable Definitions

Please see the Codebook for information regarding variables in this segment.

10.3.35.3 Special Notes

This questionnaire is administered the summer following the year of interest. The RXQ questions for the reference year 2022 were asked in the summer of 2023. Special non-response adjustment weights are included in the file to account for survey non-response from the fall to the summer data collection period. Note that counts of cases with positive Topical weights may vary within the data year and may change across years due to response rates, sample sizes, and fielding methods. The Topical weights account for these changes. Please see section 9.4.2 for information on using Topical weights.

The "other specify" questions PDNOOS and PDNTOS are back coded as necessary into the reason(s) for not using the current coverage response options and the reason(s) for not being enrolled response options, respectively, but the verbatim text is not released.

10.3.36 Satisfaction with Care (SATWCARE)

10.3.36.1 Core Content

The Satisfaction with Care segment contains data from the SCQ section on satisfaction with different aspects of medical care, such as cost and the information provided by the beneficiary's medical care provider. The

questions about satisfaction with care represent the respondent's general opinion of all medical care received in the year preceding the interview.

10.3.36.2 Variable Definitions

Please see the Codebook for information regarding variables in this segment.

10.3.36.3 Special Notes

N/A

10.3.37 Social Determinants of Health (SDOH)

10.3.37.1 Core Content

The Social Determinants of Health segment contains information on beneficiaries' residential history and geographic-level social determinants of health.

10.3.37.2 Variable Definitions

STATE: State of residence as of 12/31.

COUNTY: FIPS county code of residence as of 12/31.

ZIP: Postal zip code of residence as of 12/31.

ADDRCHNG: Flag for change in address since prior fall round. This is set to 1 if the beneficiary changed addresses.

DISTANCE: The distance moved in miles if the beneficiary changed addresses from one fall round to the next. This is rounded to a whole number and released as a range.

NUMADDR: Total number of unique addresses.

SDI: Social Deprivation Index (SDI) based on the percentile rank of the beneficiary's county of residence.

SVI: Social Vulnerability Index (SVI) based on the percentile rank of the beneficiary's county of residence.

HOSPBEDS: Number of hospital beds per 1,000 population in the beneficiary's county of residence.

LTCBEDS: Number of nursing facility beds per 1,000 population in the beneficiary's county of residence.

MDRATE: Total number of non-federal MDs per 1,000 population in the beneficiary's county of residence.

10.3.37.3 Special Notes

The 2022 SDOH segment only contains information from Fall 2022, meaning there is only information associated with the beneficiary's current address as of Fall 2022.

The SDI variable on the 2022 Survey File is sourced from the Social Deprivation Index (SDI) and uses 2019 county-level data.³⁰ SDI is a geographic-level SDOH index initially developed by Butler et. al in 2012 using American Community Survey (ACS) data and released by the Robert Graham Center. SDI offers a composite measure calculated from seven socio-demographic ACS variables to identify socio-economically disadvantaged communities.

The SVI variable on the 2022 Survey File is sourced from the Social Vulnerability Index (SVI) and uses 2022 county-level data. SVI is a geographic-level SDOH index initially developed by the CDC and ATSDR in 2011 using ACS data.³¹ SVI offers a composite measure calculated from 16 ACS variables on socio-economic status, household characteristics, racial and ethnic minorities, and housing and transportation to identify potentially vulnerable communities.

The HOSPBEDS, LTCBEDS, and MDRATE variables in the 2022 Survey File are sourced from AHRQ's Social Determinants of Health Database and use 2019 county-level data.³²

10.3.38 Telemedicine (TELEMED)

10.3.38.1 Core Content

The Telemedicine segment contains data from TLQ about the availability of telemedicine visits and the beneficiary's use of telemedicine visits.

10.3.38.2 Variable Definitions

Please see the Codebook for information regarding variables in this segment.

10.3.38.3 Special Notes

This questionnaire is administered the winter following the year of interest. Special non-response adjustment weights are included in the file to account for survey non-response from the fall to the winter data collection period. Note that counts of cases with positive Topical weights may vary within the data year and may change across years due to response rates, sample sizes, and fielding methods. The Topical weights account for these changes. Please see section 9.4.2 for information on using Topical weights.

10.3.39 Usual Source of Care (USCARE)

10.3.39.1 Core Content

The Usual Source of Care segment contains data from USQ on where and how the beneficiary typically seeks medical care.

10.3.39.2 Variable Definitions

Please see the Codebook for information regarding variables in this segment.

³⁰ "Social Deprivation Index (SDI)," Robert Graham Center, <https://www.graham-center.org/maps-data-tools/social-deprivation-index.html>.

³¹ "CDC/ATSDR Social Vulnerability Index (CDC/ATSDR SVI): Overview," Centers for Disease Control and Prevention and Agency for Toxic Substances and Disease Registry, <https://www.atsdr.cdc.gov/placeandhealth/svi/index.html>.

³² "Social Determinants of Health Database," Agency for Healthcare Research and Quality, <https://www.ahrq.gov/sdoh/data-analytics/sdoh-data.html>.

10.3.39.3 Special Notes

Several "other specify" variables are back coded as necessary into response options, but the verbatim text is not released. Back coded "other specify" variables include PVSPEC (provider specialty), LANGPREF (the language in which the beneficiary prefers to receive medical care), GETUSOS (how beneficiary normally gets to their provider), ACCOTHOS (why someone accompanies the beneficiary to their appointments), PLACEKND (the kind of place the beneficiary goes for medical care), and USWHYNAV (why the beneficiary's usual doctor is not available).

This questionnaire is administered the winter following the year of interest. Special non-response adjustment weights are included in the file to account for survey non-response from the fall to the winter data collection period. Note that counts of cases with positive Topical weights may vary within the data year and may change across years due to response rates, sample sizes, and fielding methods. The Topical weights account for these changes. Please see section 9.4.2 for information on using Topical weights.

10.3.40 Vision and Hearing (VISHEAR)

10.3.40.1 Core Content

The Vision and Hearing segment contains information on the beneficiary's eye health and hearing status.

10.3.40.2 Variable Definitions

Please see the Codebook for information regarding variables in this segment.

10.3.40.3 Special Notes

The "other specify" question EDOCTYOS is back coded as necessary into a variable (EDOCTYPE) capturing the type of doctor the beneficiary saw at their last eye exam, but the verbatim text is not released.

Six global disability questions are released to comply with HHS guidance. DISHEAR and DISSEE are included on the VISHEAR segment. Variables DISDECSN, DISWALK, DISBATH, and DISERRND are included on the NAGIDIS segment.

10.3.41 COVID-19 Facility Beneficiary-Level (FBENCVFL)

10.3.41.1 Core Content

The COVID-19 Facility Beneficiary-Level segment contains information collected in the CV section in Fall 2022 and Winter 2023, including COVID-19 vaccination, diagnosis, testing, and care received by different types of health care providers.

10.3.41.2 Variable Definitions

Please see the Codebook for information regarding variables in this segment.

10.3.41.3 Special Notes

This segment combines data from the Fall 2022 and Winter 2023 rounds, including recorded doses of COVID-19 vaccines. Many of the variables on the FBENCVFL segment are similar to variables available on the Survey File segments containing Community data from the COVID-19 Questionnaire.

In Winter 2023, the question text and response options at VACNME were updated to include the COVID-19 vaccine Novavax.

10.3.42 COVID-19 Facility Facility-Level (FFACCVFL)

10.3.42.1 Core Content

The COVID-19 Facility Facility-Level segment contains COVID-19 related information collected in the FC section in Fall 2022 and Winter 2023, including telehealth services provided, suspension of in-person services, prevention activities, prospective vaccination policies for staff and residents, personnel changes, mental health services provided, and social/recreational services provided.

10.3.42.2 Variable Definitions

Please see the Codebook for information regarding variables in this segment.

10.3.42.3 Special Notes

This segment combines data from the Fall 2022 and Winter 2023 rounds.

10.3.43 Weights

For information about the ever enrolled and continuously enrolled cross-sectional weights and two-year, three-year, and four-year longitudinal weights available in the Survey File LDS and obtaining weighted estimates using these files, please see section 9.4.

For discussion on how the weights files were created, please refer to the *MCBS Methodology Report*, which can be found on the CMS website at <https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS/Codebooks>.

11. REFERENCES

- Agency for Healthcare Research and Quality. "Social Determinants of Health Database."
<https://www.ahrq.gov/sdoh/data-analytics/sdoh-data.html>.
- Centers for Disease Control and Prevention and Agency for Toxic Substances and Disease Registry.
 "CDC/ATSDR Social Vulnerability Index (CDC/ATSDR SVI): Overview."
<https://www.atsdr.cdc.gov/placeandhealth/svi/index.html>.
- Centers for Medicare & Medicaid Services. "Long-Term Care Facility Resident Assessment Instrument 3.0
 User's Manual v1.17.1." October 2019.
- 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.
- Kalton, Graham, and Daniel Kasprzyk. "The Treatment of Missing Survey Data." *Survey Methodology* 12, no. 1
 (1986): 1-16.
- National Center for Health Statistics. "2022 National Health Interview Survey (NHIS) Questionnaire." Centers
 for Disease Control and Prevention. 2022.
https://ftp.cdc.gov/pub/Health_Statistics/NCHS/Survey_Questionnaires/NHIS/2022/EnglishQuest-508.pdf
- The Robert Graham Center. "Social Deprivation Index (SDI)." <https://www.graham-center.org/maps-data-tools/social-deprivation-index.html>.
- University of Wisconsin School of Medicine Public Health. "2020 Area Deprivation Index v3.2."
<https://www.neighborhoodatlas.medicine.wisc.edu>
- "U.S. FDA Grants Emergency Use Authorization for Novavax COVID-19 Vaccine, Adjuvanted for Individuals
 Aged 18 and Over." Novavax Inc., 2022. <https://ir.novavax.com/press-releases/2022-07-13-U-S-FDA-Grants-Emergency-Use-Authorization-for-Novavax-COVID-19-Vaccine%2C-Adjuvanted-for-Individuals-Aged-18-and-Over>
- Wolter, Kirk. *Introduction to variance estimation*. Springer Science & Business Media, 2007.

APPENDICES

12. APPENDICES

Appendix A: MCBS Rounds by Data Year and Season

Year	Winter	Summer	Fall
1991	n/a	n/a	1
1992	2	3	4
1993	5	6	7
1994	8	9	10
1995	11	12	13
1996	14	15	16
1997	17	18	19
1998	20	21	22
1999	23	24	25
2000	26	27	28
2001	29	30	31
2002	32	33	34
2003	35	36	37
2004	38	39	40
2005	41	42	43
2006	44	45	46
2007	47	48	49
2008	50	51	52
2009	53	54	55
2010	56	57	58
2011	59	60	61
2012	62	63	64
2013	65	66	67
2014	68	69	70
2015	71/72	71/72	73
2016	74	75	76
2017	77	78	79
2018	80	81	82
2019	83	84	85
2020	86	87	88
2021	89	90	91
2022	92	93	94
2023	95	96	97
2024	98	99	100

Appendix B: Sample Code³³

Joining Segments within the 2022 Survey File LDS

Data users can join segments within and/or across the Survey File and Cost Supplement File. What follows below is a hypothetical research question with sample SAS code for the construction of an analytic file. In this example, the MCBS is interested in studying the self-reported general health for Medicare beneficiaries living in the community with diabetes.

First, there are two measures required to identify our study population: residence status and self-reported diabetes. Variables corresponding to these measures can be found in the following Survey File segments, respectively: Demographics (DEMO) and Chronic Conditions (CHRNCOND). General health information is found in the General Health (GENHLTH). To ensure estimates are representative of the continuously enrolled Medicare population, the MCBS will also require weights from the CENWGTS file. Please note that if using an MCBS Survey File LDS Topical segment (such as ACCSSMED, CHRNPAIN, etc.), users should instead join onto the Topical segment and use the adjusted weights included with that segment.

Below, is an example of how multiple Survey File segments can be joined with the CENWGTS segment in SAS using BASEID as the key variable. Sample code is provided for two alternative methods for joining MCBS data, one using PROC SQL and one using SAS merge; users can use their preferred method. When joining segments, all observations in the CENWGTS segment should be preserved.

SAS (PROC SQL Join Method)

```
PROC SQL;
CREATE TABLE joined AS
  SELECT A.*,
         B.h_age,
         B.int_type,
         C.d_ocdtyp,
         D.genhelth
  FROM surveyyy.cenwgts AS A
  LEFT JOIN surveyyy.demo AS B
    ON A.baseid = B.baseid
  LEFT JOIN surveyyy.chrncond AS C
    ON A.baseid = C.baseid
  LEFT JOIN surveyyy.genhlth AS D
    ON A.baseid = D.baseid;
QUIT;
```

SAS (SAS Merge Method)

```
Data merged;
  merge surveyYY.CENWGTS (in = a)
        surveyYY.DEMO (keep = BASEID H_AGE INT_TYPE)
        surveyYY.CHRNCOND (keep = BASEID D_OCDTYP)
        surveyYY.GENHLTH (keep = BASEID GENHELTH);
  by BASEID;
  if a;
run;
```

³³ The "YY" in "costYY" and "surveyYY" refers to the data year of the Cost Supplement File and Survey File, respectively. Longitudinal code is represented with the convention of Y1, Y2, etc.

In order to segment the file to beneficiaries living in the community only, subset the file on the variable INT_TYPE.

```
Data joined_surveyfile;
    set joined;
    where INT_TYPE = 'C'; /* denotes individuals living only in the community */
run;
```

Now there is an analytic file that includes all the Survey File variables and weights required to analyze general health for Medicare beneficiaries living in the community with diabetes. Data users can export the created dataset for use with R and Stata.

Repeated Cross-Sectional or Pooled Analysis (Section 9.7.2)

Sample code

The sample code below demonstrates the steps involved in constructing a repeated cross-sectional or pooled analytic dataset and performing analysis. The example below estimates percent of Medicare beneficiaries that are dually eligible (i.e., enrolled in both Medicare and Medicaid programs) during the prior data year and the current data year.

Although the MCBS includes variables to obtain weighted estimates and estimated standard errors using Taylor-series linearization approach, the balanced repeated replication (BRR) method, also known as Fay's method, provides more analytic flexibility when performing analysis using pooled cross-sectional data.³⁴

CMS generally recommends the BRR method of variance estimation to MCBS users because it requires neither the specification of strata and cluster definitions nor the specification of domain or subgroup definitions in subpopulation analyses, which are required for Taylor-series estimation and are common inadvertent omissions. However, the Taylor series method of variance estimation is also appropriate for experienced users who prefer this method or in instances where the BRR method is not possible in the available software. For these reasons, the MCBS data files include the variables SUDSTRAT and SUDUNIT, which are needed for Taylor-series estimation. The SAS functions %surveyglm and %surveygenmod appropriately allow for strata and cluster definitions. When using these functions (and in any other instances where Taylor series estimation is used), specify SUDSTRAT as the strata definitions and SUDUNIT as the cluster definitions.

The examples presented in this section involve multiple years of MCBS data and use replicate weights – a form of the BRR technique.

Example

/* Join prior data year administrative records (HISUMRY) file with cross-sectional weights (CENWGTS) file */

SAS (PROC SQL Join Method)

```
PROC SQL;
CREATE TABLE mcbsY1 AS
SELECT A.*,
       B.h_opmdcd
FROM   surveyY1.cenwgts(DROP=version) AS A
LEFT JOIN surveyY1.hisumry AS B
```

³⁴ Given the rotating panel design of the MCBS, performing pooled cross-sectional analysis using Taylor-Series Linearization method of variance estimation will require additional adjustments to account for non-independence of beneficiaries across years in a multi-year dataset.


```
ON A.baseid = B.baseid;
```

```
QUIT;
```

SAS (SAS Merge Method)

```
data mcbsY1;
merge surveyY1.CENWGTS (in = a drop = VERSION)
      surveyY1.HISUMRY (keep = BASEID H_OPMDCD);
  by BASEID;
  if a;
run;
```

```
/* Create Analytic Dataset for Repeated Cross-Sectional or Pooled Analysis */
```

```
/* Join current data year administrative records (HISUMRY) file with cross-sectional weights (CENWGTS) file */
```

SAS (PROC SQL Join Method)

```
PROC SQL;
CREATE TABLE mcbsY2 AS
  SELECT A.*,
         B.h_opmdcd
  FROM   surveyY2.cenwgts(DROP=version) AS A
  LEFT JOIN surveyY2.hisumry AS B
        ON A.baseid = B.baseid;
QUIT;
```

SAS (SAS Merge Method)

```
data mcbsY2;
  merge surveyY2.CENWGTS (in = a drop = VERSION)
        surveyY2.HISUMRY (keep = BASEID H_OPMDCD);
  by BASEID;
  if a;
run;
/* Concatenate prior and current cross-sectional files */
data mcbs_analytic_file;
  set mcbsY1 mcbsY2;
run;
```

SAS

* Estimate Percent of Dually Eligible Medicare Beneficiaries (Pooled estimate representing the moving average of nationally representative year-specific estimates) using balanced repeated replication (Fay's method));

```
proc surveyfreq data = mcbs_analytic_file varmethod = brr (fay=.30);
  table H_OPMDCD;
  weight CEYRSWGT;
  repweights CEYRS001-CEYRS100;
run;
```

* Estimate Percent of Dually Eligible Medicare Beneficiaries by Year (nationally representative, year-specific estimates) using balanced repeated replication (Fay's method);

```
proc surveyfreq data = mcbs_analytic_file varmethod = brr (fay=.30);
  table SURVEYYR * H_OPMDCD/ row;
  weight CEYRSWGT;
```

```
repweights CEYRS001-CEYRS100;
run;
```

Stata

```
* Declare survey dataset
svyset _n [pweight = CEYRSWGT], brrweight(CEYRS001-CEYRS100) fay(.3) vce(brr)

* Estimate Percent of Dually Eligible Medicare Beneficiaries (Pooled estimate representing the
* moving average of nationally representative year-specific estimates)
svy brr, fay(.3): tab H_OPMDCD

* Estimate Percent of Dually Eligible Medicare Beneficiaries (nationally representative, year-specific estimates)
svy brr, fay(.3): tab H_OPMDCD SURVEYR, column
```

R

Note: Data users will need to install the 'survey' package to use the svrepdesign function below.

```
# Specify survey design object
mcbs <- svrepdesign(
  weights = ~CEYRSWGT,
  repweights = "CEYRS[001-100]+",
  type = "Fay",
  rho = 0.3,
  data = mcbs_analytic_file,
  combined.weights = TRUE
)

# Estimate Percent of Dually Eligible Medicare Beneficiaries by Year (Pooled estimate representing the moving
average of nationally representative year-specific estimates)
prop.table(svytable(~H_OPMDCD, design=mcbs))

# Estimate Percent of Dually Eligible Medicare Beneficiaries by Year (nationally representative, year-specific
estimates)
prop.table(svytable(~H_OPMDCD + SURVEYR, design=mcbs), 2)
```

Conducting Subgroup Analyses with Appropriate Variance Estimation

Using the BRR method of variance estimation

Variance estimation can be impacted by selecting individuals prior to analysis. If the BRR variance estimation method is used, subgroup analyses can be conducted by limiting the dataset to the desired sub-sample. There are multiple ways to conduct subgroup analyses using BRR.

For indicator variables in three-way tables, you can create flags to help you identify the population of interest. For instance, if you are interested in the prevalence of diabetes in men versus women, but only in the over-65 population in Medicare Advantage, you could use the following SAS code:

```
proc surveyfreq data=mcbsdata VARMETHOD = brr (fay=.30);
  table SEX * DIABETES * OVER65MA / col notot;
  weight CEYRSWGT;
  repweights CEYRS001-CEYRS100;
run;
```

This sample code assumes an analytic data set, including replicate weights, in which the data user has created binary analytic variables for SEX and DIABETES, as well as a flag variable, OVER65MA, to identify the population of interest for this analysis. In this case, the flag is equal to 1 if the beneficiary is over 65 and in Medicare Advantage, and equal to 0 otherwise.

Since variance estimation using the BRR approach permits limiting the dataset to the desired sub-sample of interest, the following SAS code can also be used to achieve the same result (shown in the frequency columns) through subgroup analysis:

```
data mcbsdata_subset;
    set mcbsdata;
    if OVER65MA = 1 then output;
run;

proc surveyfreq data=mcbsdata_subset VARMETHOD = brr (fay=.30);
    table SEX * DIABETES / col notot nopercnt;
    weight CEYRSWGT;
    repweight CEYRS001-CEYRS100;
run;
```

Using the Taylor Series linearization method of variance estimation

If other variance estimation methods, such as Taylor Series linearization are used, the correct way to analyze MCBS data is to employ domain statements (in SAS: proc surveymeans, surveylogistic, and surveyreg) or indicator variables in three-way tables (in SAS: proc surveyfreq). 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.

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

```
proc surveyfreq data=mcbsdata;
    table SEX * DIABETES * OVER65MA / col notot;
    strata SUDSTRAT;
    cluster SUDUNIT;
    weight CEYRSWGT;
run;
```

Appendix C: Initial Interview Variables

Exhibit C.1: Initial Interview Variables

Segment	Topic	LDS Variable Name
DEMO	Date of Birth	D_DOB
DEMO	Sex	ROSTSEX
DEMO	Hispanic Origin	HISPORIG HISPORMA HISPORPR HISPORCU HISPOROT
DEMO	Race	D_RACE2 RACEAA RACEAS RACENH RACEWH RACEAI
DEMO	Asian Race Subcategories	RACEASAI RACEASCH RACEASFI RACEASJA RACEASKO RACEASVI RACEASOT
DEMO	Pacific Islander Race Subcategories	RACEPIHA RACEPIGU RACEPISA RACEPIOT
DEMO	Military Service	SPAFEVER SPAFVIET SPAFKORE SPAFWWII SPAFGULF SPAFIRAF SPAFPEAC SPNGEVER SPNGALL SPNGDSBL SPVARATE
DEMO	Number of Children	SPCHNLNM
DEMO	Limited English Proficiency	ENGWELL ENGREAD OTHLANG WHATLANG
DEMO	Education	SPDEGRCV

Segment	Topic	LDS Variable Name
DEMO	Income	INCOME
CHRNCOND	Reason for Medicare Eligibility	EMHBP
		EMMYOCAR
		EMCHD
		EMCFAIL
		EMHRTCND
		EMSTROKE
		EMCSKIN
		EMCANCER
		EMARTERY
		EMARTHHR
		EMARTOST
		EMARTHOT
		EMMENTAL
		EMALZMER
		EMDEMENT
		EMDEPRSS
		EMPSYCHO
		EMOSTEOP
		EMBRKHIP
		EMPARKIN
		EMEMPHYS
		EMPPARAL
		EMAMPUTE
		EMDIABTS
		EMOTHOS
CHRNCOND	Number of Medications Taken for Blood Pressure	HYPEMANY
FACCHAR	Place of Residence before Facility Admission	BEFORADM
FACCHAR	Household Makeup before Facility Admission	D_LIVWTH

Appendix D: Table of Links to MCBS Documentation

MCBS Resources	Links
CMS MCBS website	https://www.cms.gov/data-research/research/medicare-current-beneficiary-survey
MCBS LDS file information	https://www.cms.gov/data-research/files-for-order/data-disclosures-and-data-use-agreements-duas/limited-data-set-lds
MCBS Microdata PUFs	https://www.cms.gov/data-research/statistics-trends-and-reports/mcbs-public-use-file
CMS Chronic Conditions Warehouse (CCW)	https://www.ccwdata.org/web/guest/home/
Data User's Guides, Methodology Reports, Codebooks, and LDS Variable Crosswalks	https://www.cms.gov/data-research/research/medicare-current-beneficiary-survey/data-documentation-codebooks
PUF Table Packages and Chartbook PDFs	https://www.cms.gov/data-research/research/medicare-current-beneficiary-survey/data-tables
Early Look, Data Briefs, Infographics, and Tutorials	https://www.cms.gov/data-research/research/medicare-current-beneficiary-survey/data-briefs-tutorials
Bibliography	https://www.cms.gov/data-research/research/medicare-current-beneficiary-survey/bibliography
Questionnaires and Questionnaire User Documentation	https://www.cms.gov/data-research/research/medicare-current-beneficiary-survey/questionnaires
MCBS Interactives – COVID-19 Data Tool, Survey File PUF Data Tool, and Financial Well-being Data Tool ³⁵	https://mcbs-interactives.norc.org/
Chartbook ³⁶	https://data.cms.gov/medicare-current-beneficiary-survey-mcbs

³⁵ The MCBS Interactives consist of three data tools, the Financial Well-being Data Tool, the Survey File PUF Data Tool and the COVID-19 Data Tool. Each tool contains multiple interactive dashboards that allow users to sort and visualize data according to a variety of demographic and health-related factors.

³⁶ Beginning with the release of 2021 MCBS data, the MCBS Chartbook website replaced the PDF version of the MCBS Chartbook that was updated and released annually on the CMS MCBS website to disseminate current estimates on the Medicare population. MCBS estimates from 2015 through 2020 can be found in both the online version of the MCBS Chartbook and the previous MCBS Chartbook PDFs at <https://www.cms.gov/data-research/research/medicare-current-beneficiary-survey/data-tables>.

Appendix E: 2022 MCBS Survey File Segments and Historic RIC Segments

Survey File Segment	Segment Abbrev	Historic RIC Segment
Access to Care	ACCESSCR	3
Access to Care, Medical Appointments	ACCSSMED	3
Administrative Utilization Summary	ADMNUTLS	A
Assistance	ASSIST	2H
Chronic Conditions	CHRNCOND	2, 2P
Chronic Conditions Flags	CHRNCDFL	N/A
Chronic Pain	CHRNPAIN	N/A
Cognitive Measures	COGNFUNC	N/A
COVID-19 Topical	COVIDTOP	N/A
COVID-19 Experiences	COVIDEXP	N/A
Demographics	DEMO	1, 9, A, K
Diabetes	DIABETES	N/A
Facility Assessments	FACASMNT	2F
Facility Characteristics	FACCHAR	7, 7S
Falls	FALLS	2, 2P
Food Insecurity	FOODINS	N/A
General Health	GENHLTH	2
Health Insurance Summary	HISUMRY	4, A
Health Insurance Timeline	HITLINE	4, A
Household (HH) Characteristics	HHCHAR	5
Income and Assets	INCASSET	1, Income Asset
Interview Characteristics	INTERV	4, 8, 9, K
Medicare Advantage (MA) Plan Questions	MAPLANQX	H
Medicare Plan Beneficiary Knowledge	MCREPLNQ	KN
Minimum Data Set	MDS3	MDS, 10
Mental Health	MENTHLTH	N/A
Mobility	MOBILITY	N/A
Multiple Year Enrollment	MYENROLL	N/A
Nagi Disability	NAGIDIS	2, 2H, 2P
Nicotine and Alcohol	NICOALCO	2, 2P
Outcome and Assessment Information	OASIS	OAS, 10
Patient Activation	PNTACT	PA
Preventive Care	PREVCARE	2, 2P
RX Medications	RXMED	RX
Residence Timeline	RESTMLN	6, 9, A, K
Satisfaction with Care	SATWCARE	3
Social Determinants of Health	SDOH	N/A
Telemedicine	TELEMED	N/A

Survey File Segment	Segment Abbrev	Historic RIC Segment
Usual Source of Care	USCARE	2, 3
Vision and Hearing	VISHEAR	2
Weights	CENWGTS	X, XE, X3, X4
	EVRWGTS	
	LNG2WGTS	
	LNG3WGTS	
	LNG4WGTS	
COVID-19 Facility Beneficiary-Level	FBENCVFL	N/A
COVID-19 Facility Facility-Level	FFACCVFL	N/A
Fee-for-Service Claims	FFS	Research Claims