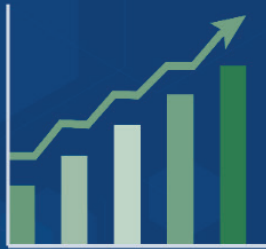


# 2018 | DATA USER'S GUIDE: SURVEY FILE



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

## Version Control Log

Date	Version	Revisions
6/1/2020	1.0	Initial version published.
9/24/2020	1.1	Revised glossary definition for exit interview.

# TABLE OF CONTENTS

<b>ACRONYM LIST .....</b>	<b>.xi</b>
<b>1. INTRODUCTION.....</b>	<b>1</b>
1.1 Contents of the Data User's Guide: Survey File.....	2
<b>2. GENERAL GUIDELINES FOR DATA USE.....</b>	<b>3</b>
2.1 Data Access.....	3
2.2 Guidelines for Citation of Data Source .....	3
<b>3. WHAT'S NEW FOR DATA YEAR 2018? .....</b>	<b>5</b>
3.1 Sampling .....	5
3.2 Questionnaires.....	5
3.2.1 General .....	5
3.2.2 Section-Specific Changes .....	6
3.3 Data Collection .....	7
3.4 Documentation .....	7
3.5 Data Processing .....	7
<b>4. SURVEY OVERVIEW .....</b>	<b>12</b>
4.1 Design of MCBS .....	12
4.2 Sample Design.....	13
4.3 Case Types.....	14
4.3.1 Incoming and Continuing Cases .....	14
4.3.2 Community Interviews and Facility Interviews.....	14
4.4 Interviewing and Training Procedures.....	15
4.4.1 Overview of Data Collection .....	15
4.4.2 Interviewer Training.....	15
4.4.3 Privacy and Data Security .....	15
4.5 Completed Interviews.....	16
4.6 Item Non-Response .....	16
<b>5. QUESTIONNAIRES .....</b>	<b>17</b>
5.1 Overview .....	17
5.2 Community Questionnaire.....	18
5.2.1 Baseline Interview .....	19

5.2.2 Continuing Interview .....	20
5.2.3 Core Questionnaire Sections .....	22
5.2.4 Topical Questionnaire Sections .....	23
5.2.5 Community Questionnaire Section Rotation within a Data Year .....	23
5.3 Facility Instrument .....	24
5.3.1 Facility Baseline Interview .....	25
5.3.2 Facility Continuing Interview .....	26
5.3.3 Facility Continuing Core Sections .....	28
<b>6. SAMPLING .....</b>	<b>29</b>
6.1 Medicare Population Covered by the 2018 MCBS Data .....	29
6.2 Targeted Population and Sampling Strata .....	33
6.3 Primary and Secondary Sampling Units .....	34
6.3.1 Eligibility: Medicare Population Covered by the 2018 LDS .....	35
6.4 Sample Selection .....	35
<b>7. DATA PRODUCTS &amp; DOCUMENTATION .....</b>	<b>37</b>
7.1 Contents of Data Release .....	37
7.1.1 2018 MCBS Survey File .....	38
7.1.2 2018 MCBS Cost Supplement File .....	43
7.1.3 Using the Data .....	45
7.1.4 Research Claims Files .....	45
7.2 Which File Do I Need? .....	46
7.2.1 Survey File Only .....	46
7.2.2 Using Both Survey File and Cost Supplement File .....	46
7.2.3 Using Both Community and Facility Data .....	47
<b>8. FILE STRUCTURE .....</b>	<b>48</b>
8.1 LDS Specifications .....	48
8.2 File Structure .....	48
8.3 Claims Files .....	48
8.3.1 Utilization Detail Records .....	48
<b>9. DATA FILE DOCUMENTATION .....</b>	<b>50</b>
9.1 LDS Contents .....	50
9.2 LDS Components .....	50
9.2.1 Codebooks .....	50
9.2.2 Questionnaires .....	51
9.2.3 Data User Resources .....	51

9.3 Data Edits and Imputation .....	51
9.3.1 Data Edits .....	51
9.3.2 Imputation .....	52
9.4 Weighting.....	53
9.4.1 Preparing Statistics (Using the Full Sample Weights) .....	53
9.5 Using the Data.....	59
9.5.1 Merging Segments within 2018 .....	59
9.6 Variance Estimation (Using the Replicate Weights) .....	59
9.6.1 Variables Available for Variance Estimation .....	59
9.6.2 Variance Estimation for Analyses of Single Year of MCBS .....	60
9.6.3 Subgroup Analysis.....	60
9.7 Combining Multiple Years of Data.....	60
9.7.1 Longitudinal Analysis.....	60
9.7.2 Repeated Cross-Sectional or Pooled Analysis.....	62
<b>10. DATA FILE NOTES .....</b>	<b>63</b>
10.1 Global Information .....	63
10.1.1 BASEID .....	63
10.1.2 Missing Values .....	63
10.1.3 Derived and Administrative Variables.....	63
10.1.4 Initial Interview Variables .....	63
10.1.5 Ever Variables.....	63
10.1.6 Data Editing.....	64
10.1.7 Other Specify Questions .....	64
10.1.8 Consistency with Medicare Program Statistics.....	64
10.1.9 Do administrative data override survey-reported data? .....	65
10.2 Survey File Segment Information.....	65
10.3 Access to Care (ACCESSCR) .....	65
10.3.1 Core Content .....	65
10.3.2 Variable Definitions .....	65
10.3.3 Special Notes.....	65
10.4 Access to Care, Medical Appointment (ACCSSMED).....	66
10.4.1 Core Content .....	66
10.4.2 Variable Definitions .....	66
10.4.3 Special Notes.....	66
10.5 Administrative Utilization Summary (ADMNUTLS).....	67
10.5.1 Core Content .....	67
10.5.2 Variable Definitions .....	67
10.5.3 Special Notes.....	73

10.6 Assistance (ASSIST) .....	73
10.6.1 Core Content .....	73
10.6.2 Variable Definitions .....	74
10.6.3 Special Notes .....	74
10.7 Chronic Conditions (CHRNCOND) .....	74
10.7.1 Core Content .....	74
10.7.2 Variable Definitions .....	74
10.7.3 Special Notes .....	75
10.8 Chronic Condition Flags (CHRNCDL) .....	75
10.8.1 Core Content .....	75
10.8.2 Variable Definitions .....	76
10.8.3 Special Notes .....	76
10.9 Chronic Pain (CHRNPAIN) .....	76
10.9.1 Core Content .....	76
10.9.2 Variable Definitions .....	76
10.9.3 Special Notes .....	77
10.10 Demographics (DEMO) .....	77
10.10.1 Core Content .....	77
10.10.2 Variable Definitions .....	77
10.10.3 Special Notes .....	82
10.11 Diabetes (DIABETES) .....	82
10.11.1 Core Content .....	82
10.11.2 Variable Definitions .....	83
10.11.3 Special Notes .....	83
10.12 Facility Assessments (FACASMNT) .....	83
10.12.1 Core Content .....	83
10.12.2 Variable Definitions .....	83
10.12.3 Special Notes .....	85
10.13 Facility Characteristics (FACCHAR) .....	86
10.13.1 Core Content .....	86
10.13.2 Variable Definitions .....	86
10.13.3 Special Notes .....	86
10.14 Falls (FALLS) .....	87
10.14.1 Core Content .....	87
10.14.2 Variable Definitions .....	87
10.14.3 Special Notes .....	87
10.15 Food Insecurity (FOODINS) .....	87
10.15.1 Core Content .....	87
10.15.2 Variable Definitions .....	87

10.15.3 Special Notes .....	87
10.16 General Health (GENHLTH) .....	87
10.16.1 Core Content .....	87
10.16.2 Variable Definitions.....	87
10.16.3 Special Notes .....	88
10.17 Health Insurance Summary (HISUMRY) .....	88
10.17.1 Core Content .....	88
10.17.2 Variable Definitions.....	88
10.17.3 Special Notes .....	92
10.18 Health Insurance Timeline (HITLINE) .....	95
10.18.1 Core Content .....	95
10.18.2 Variable Definitions.....	95
10.18.3 Special Notes .....	96
10.19 Household Characteristics (HHCHAR) .....	97
10.19.1 Core Content .....	97
10.19.2 Variable Definitions.....	97
10.19.3 Special Notes .....	98
10.20 Income and Assets (INCASSET).....	98
10.20.1 Core Content .....	98
10.20.2 Variable Definitions.....	98
10.20.3 Special Notes .....	98
10.21 Interview Characteristics (INTERV) .....	99
10.21.1 Core Content .....	99
10.21.2 Variable Definitions.....	99
10.21.3 Special Notes .....	100
10.22 Medicare Advantage Plan Questions (MAPLANQX).....	100
10.22.1 Core Content .....	100
10.22.2 Variable Definitions.....	100
10.22.3 Special Notes .....	100
10.23 Medicare Plan Beneficiary Knowledge (MCREPLNQ).....	100
10.23.1 Core Content .....	100
10.23.2 Variable Definitions.....	101
10.23.3 Special Notes .....	101
10.24 Mental Health (MENTHLTH) .....	101
10.24.1 Core Content .....	101
10.24.2 Variable Definitions.....	101
10.24.3 Special Notes .....	101

10.25 Mobility (MOBILITY) .....	101
10.25.1 Core Content .....	101
10.25.2 Variable Definitions.....	101
10.25.3 Special Notes .....	102
10.26 Minimum Data Set (MDS3) .....	102
10.26.1 Core Content .....	102
10.26.2 Variable Definitions.....	102
10.26.3 Special Notes .....	102
10.27 NAGI Disability (NAGIDIS) .....	102
10.27.1 Core Content .....	102
10.27.2 Variable Definitions.....	102
10.27.3 Special Notes .....	104
10.28 Nicotine and Alcohol (NICOALCO) .....	104
10.28.1 Core Content .....	104
10.28.2 Variable Definitions.....	104
10.28.3 Special Notes .....	104
10.29 Outcome and Assessment Information (OASIS) .....	105
10.29.1 Core Content .....	105
10.29.2 Variable Definitions.....	105
10.29.3 Special Notes .....	105
10.30 Patient Activation (PNTACT) .....	105
10.30.1 Core Content .....	105
10.30.2 Variable Definitions.....	105
10.30.3 Special Notes .....	105
10.31 Preventive Care (PREVCARE).....	105
10.31.1 Core Content .....	105
10.31.2 Variable Definitions.....	105
10.31.3 Special Notes .....	105
10.32 Residence Timeline (RESTMLN) .....	106
10.32.1 Core Content .....	106
10.32.2 Variable Definitions.....	106
10.32.3 Special Notes .....	106
10.33 RX Medications (RXMED) .....	107
10.33.1 Core Content .....	107
10.33.2 Variable Definitions.....	107
10.33.3 Special Notes .....	107
10.34 Satisfaction with Care (SATWCARE) .....	107
10.34.1 Core Content .....	107
10.34.2 Variable Definitions.....	107

10.34.3 Special Notes .....	108
10.35 Usual Source of Care (USCARE) .....	108
10.35.1 Core Content .....	108
10.35.2 Variable Definitions.....	108
10.35.3 Special Notes .....	108
10.36 Vision and Hearing (VISHEAR).....	108
10.36.1 Core Content .....	108
10.36.2 Variable Definitions.....	108
10.36.3 Special Notes .....	108
10.37 Weights .....	109
10.37.1 Cross Sectional Weights.....	109
10.37.2 Longitudinal Weights .....	109
10.37.3 Using weights with data from Topical Questionnaire Sections .....	110
<b>11. REFERENCES.....</b>	<b>111</b>
<b>12. APPENDICES.....</b>	<b>113</b>
Appendix A: MCBS Common Definitions.....	113
Appendix B: MCBS Rounds by Data Year and Season .....	118
Appendix C: Sample Code .....	119
Appendix D: Initial Interview Variables .....	123
Appendix E: Table of Links to MCBS Documentation.....	125

## LIST OF EXHIBITS

<b>Exhibit 3.5.1:</b>	<b>2018 MCBS Content Additions.....</b>	<b>8</b>
<b>Exhibit 4.1.1:</b>	<b>Typical MCBS Data Collection Year.....</b>	<b>12</b>
<b>Exhibit 4.1.2:</b>	<b>MCBS Beneficiary Participation Timeline .....</b>	<b>13</b>
<b>Exhibit 4.2:</b>	<b>2014-2018 MCBS Rotating Panel Design .....</b>	<b>14</b>
<b>Exhibit 4.5:</b>	<b>2018 MCBS Fall Round Completed Interviews: Continuing and Incoming Panels .....</b>	<b>16</b>
<b>Exhibit 5.1:</b>	<b>MCBS Questionnaire Overview .....</b>	<b>17</b>
<b>Exhibit 5.2:</b>	<b>Overview of the MCBS Community Questionnaire Components .....</b>	<b>19</b>
<b>Exhibit 5.2.1:</b>	<b>2018 Data Collection Year MCBS Community Questionnaire Flow for Baseline Interview.....</b>	<b>20</b>
<b>Exhibit 5.2.2:</b>	<b>2018 Data Collection Year MCBS Community Questionnaire Flow for Continuing Interview.....</b>	<b>21</b>
<b>Exhibit 5.2.3:</b>	<b>2018 Data Year MCBS Community Core Sections by Data File and Data Collection Schedule* .....</b>	<b>22</b>
<b>Exhibit 5.2.4:</b>	<b>2018 Data Year MCBS Community Topical Sections by Data File and Data Collection Schedule .....</b>	<b>23</b>
<b>Exhibit 5.2.5:</b>	<b>2018 Data Collection Year MCBS Community Questionnaire Section Rotation .....</b>	<b>24</b>
<b>Exhibit 5.3:</b>	<b>Overview of the MCBS Facility Instrument .....</b>	<b>25</b>
<b>Exhibit 5.3.1:</b>	<b>2018 Data Collection Year MCBS Facility Instrument Flow for Baseline Interview .....</b>	<b>26</b>
<b>Exhibit 5.3.2:</b>	<b>2018 Data Collection Year MCBS Facility Instrument Flow for Continuing Interviews.....</b>	<b>27</b>
<b>Exhibit 5.3.3:</b>	<b>2018 Data Year MCBS Facility Core Sections by Data File and Data Collection Schedule* .....</b>	<b>28</b>
<b>Exhibit 6.1:</b>	<b>2018 MCBS Composition of Panels in LDS Data Files.....</b>	<b>29</b>

<b>Exhibit 6.1.1:</b>	<b>2018 Total Estimated Number of Medicare Beneficiaries by Sex and Race.....</b>	<b>30</b>
<b>Exhibit 6.1.2:</b>	<b>2018 Estimated Number of Male Medicare Beneficiaries by Race and Age .....</b>	<b>31</b>
<b>Exhibit 6.1.3:</b>	<b>2018 Estimated Number of Female Medicare Beneficiaries by Race and Age ..</b>	<b>32</b>
<b>Exhibit 6.2.1:</b>	<b>2018 MCBS Sampling Strata.....</b>	<b>33</b>
<b>Exhibit 6.2.2:</b>	<b>2018 Panel of Selected Beneficiaries by Hispanic and Non-Hispanic Ethnicity Classification and Age Category .....</b>	<b>34</b>
<b>Exhibit 6.4:</b>	<b>2018 MCBS Sample Selection for the LDS Releases.....</b>	<b>36</b>
<b>Exhibit 7.1:</b>	<b>2018 Contents of Data Releases .....</b>	<b>38</b>
<b>Exhibit 7.1.1:</b>	<b>2018 MCBS Survey File Segments and Contents.....</b>	<b>39</b>
<b>Exhibit 7.1.2:</b>	<b>2018 MCBS Cost Supplement File Segments and Contents .....</b>	<b>44</b>
<b>Exhibit 7.2:</b>	<b>MCBS Populations in Data Products .....</b>	<b>46</b>
<b>Exhibit 9.3.1:</b>	<b>Data Review and Editing Codes.....</b>	<b>52</b>
<b>Exhibit 9.4.1:</b>	<b>2018 MCBS Data Files Summary of Weights .....</b>	<b>55</b>
<b>Exhibit 9.7.1:</b>	<b>Constructing a Longitudinal Analytic File.....</b>	<b>61</b>
<b>Exhibit 9.7.2:</b>	<b>Constructing a Repeated Cross-Section or Pooled Analytic File .....</b>	<b>62</b>
<b>Exhibit 10.10.2a:</b>	<b>Primary RUCA (H_RUCA1) Codes, 2010 .....</b>	<b>79</b>
<b>Exhibit 10.10.2b:</b>	<b>Secondary RUCA (H_RUCA2) Codes, 2010 .....</b>	<b>80</b>
<b>Exhibit 10.12.1:</b>	<b>LDS Segments with Similar Topics for Community and Facility Interviews .....</b>	<b>85</b>
<b>Exhibit 10.12.2:</b>	<b>Differences between FACASMNT and MDS3 Data .....</b>	<b>86</b>
<b>Exhibit 10.17.1:</b>	<b>Mapping of H_PDLS01-12 Values from 2013 and Prior to 2015 and Beyond ...</b>	<b>91</b>
<b>Exhibit D.1:</b>	<b>Initial Interview Variables.....</b>	<b>123</b>

## ACRONYM LIST

ACCESSCR	Access to Care segment
ACCSSMED	Access to Care, Medical Appointment segment
ACO	Accountable Care Organization
ACQ	Access to Care Questionnaire
ACS	American Community Survey
ADLs	Activities of Daily Living
ADMNUTLS	Administrative Utilization Summary segment
AGREESCL	Agreement Scale
ASSIST	Assistance segment
ATC	Access to Care
AVQ	Address Verification Questionnaire
BQ	Background Questionnaire
BRR	Balanced repeated replication (or Fay's method)
CAPI	Computer-Assisted Personal Interviewing
CATI	Computer-Assisted Telephone Interviewing
CAU	Cost and Use
CENWGTS	Continuously enrolled weights
CHRNCOND	Chronic Conditions segment
CHRNCDFL	Chronic Condition Flags segment
CHRNPAIN	Chronic Pain segment
CMS	Centers for Medicare & Medicaid Services
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
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
DUQ	Dental Utilization Questionnaire
ENS	Enumeration Summary Questionnaire
EOBs	Explanation of Benefit Statements
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
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
HHC	Health and Health Care of the Medicare Population
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
HIS	Health Insurance Summary Questionnaire
HISUMRY	Health Insurance Summary segment
HITLINE	Health Insurance Timeline segment
HMO	Health Maintenance Organization
HS	Health Status
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)
LEP	Limited English Proficiency
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
MB	Medicare Beneficiary
MBQ	Mobility of Beneficiaries Questionnaire
MBSF	Master Beneficiary Summary File
MCBS	Medicare Current Beneficiary Survey
MCREPLNQ	Medicare Plan Beneficiary Knowledge segment
MDS	Minimum Data Set
MENTHLTH	Mental Health segment
MOBILITY	Mobility segment
MPE	Medical Provider Events segment
MPQ	Medical Provider Utilization Questionnaire
NAGIDIS	NAGI Disability segment
NHATS	National Health and Aging Trends Study
NICOALCO	Nicotine and Alcohol segment
NORC	NORC at the University of Chicago
NSQ	No-Statement Section Questionnaire
OASIS	Outcome and Assessment Information segment
OEDA	Office of Enterprise Data and Analytics

OM	Other Medical Expenses
OMB	Office of Management and Budget
OMQ	Other Medical Expenses Utilization Questionnaire
OPE	Outpatient Hospital Events segment
OPQ	Outpatient Utilization Questionnaire
PAQ	Patient Activation Questionnaire
PDP	Prescription Drug Plan
PFFS	Private Fee-for-Service
PHQ	Patient Health Questionnaire depression screening tool (PHQ-9)
PII	Personally Identifiable Information
PM	Prescription Medicine
PME	Prescribed Medicine Events segment
PMQ	Prescribed Medicine Questionnaire
PMS	Prescribed Medicine Summary
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
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
SCQ	Satisfaction with Care Questionnaire
SNF	Skilled Nursing Facility
SS	Service Summary segment
SSN	Social Security Number
SSU	Secondary Sampling Units
STQ	Statement Section 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

# 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). The MCBS is designed to aid CMS in administering, monitoring, and evaluating Medicare programs. A leading source of information on Medicare and its impact on beneficiaries, the MCBS provides important information on beneficiaries that is not available in CMS administrative data and plays an essential role in monitoring and evaluating beneficiary health status and health care policy.

The MCBS is a continuous, in-person, multi-purpose longitudinal survey, representing the population of beneficiaries aged 65 and over and beneficiaries aged 64 and below with disabilities, residing in the United States. Fieldwork for the first round of data collection began in September 1991; since then, the MCBS has continued to collect and provide essential data on the costs, use, and health care status of Medicare beneficiaries. The MCBS has conducted continuous data collection for over 25 years, completing more than one 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.

Annually, CMS releases three sets of files – a Public Use File (PUF) 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 the beneficiary 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.

**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.**

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, 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 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 2018 segments. Note that for analyses on 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 PUF, including a codebook and additional documentation, can be found at <https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/MCBS-Public-Use-File>.

For questions or suggestions on this document or other MCBS data-related questions, please email [MCBS@cms.hhs.gov](mailto: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 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.<sup>1</sup>

This document contains an overview of the survey, questionnaires, sample design, and other topics relevant to the MCBS in general. Data users can access this Guide along with other data documentation at <https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS/Codebooks>. Please also see Appendix A: MCBS Common Definitions for descriptions of frequently used or key terms.

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.
- Sections 4-9: Overview of the MCBS – These sections provide an overview of the MCBs, 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 common definitions and sample programs for data users.

---

<sup>1</sup> The Cost Supplement LDS and companion *Data User's Guide* is released three to four months after the Survey File LDS.

## 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.

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.

### 2.1 Data Access

In order to gain access to the LDS, data users must complete several steps. First, data users must sign and submit a CMS DUA and complete an LDS Worksheet. The DUA acknowledges the user's agreement to CMS' terms around data exchange, privacy, use, and storage. The LDS worksheet provides CMS with information about the research project, the specific files needed, and payment information for administrative fees associated with the data request.

Administrative processing fees for obtaining the LDS files are \$300 for the 2018 Survey File alone, and \$600 for the 2018 Survey File with the 2018 Cost Supplement (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 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).

For additional information on data access and the DUA Process, data users can visit the CMS' LDS website at [https://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/Data-Disclosures-Data-Agreements/DUA - NewLDS](https://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/Data-Disclosures-Data-Agreements/DUA-NewLDS).

Questionnaires, codebooks, and bibliographies for each survey year are available for download on the CMS' MCBS website at <https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS>. 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 document was produced, published, 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 *2018 MCBS Data User's Guide* should read:

SOURCE: Centers for Medicare & Medicaid Services. 2018 Medicare Current Beneficiary Survey Data User's Guide: Survey File. Retrieved from [ADD URL], 2018.

**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, 2018.

SOURCE: Centers for Medicare & Medicaid Services. Medicare Current Beneficiary Survey, Cost Supplement File data. Baltimore, MD: U.S. Department of Health and Human Services, 2018.

### 3. WHAT'S NEW FOR DATA YEAR 2018?

Below are the highlights and updates for the 2018 data year.

#### 3.1 Sampling

**Sample eligibility:** Beginning in 2018, data collection in Puerto Rico was officially halted.

**Hispanic oversample:** 2018 was the final year of the sampling of 75 *additional* continental U.S. Hispanics<sup>2</sup> to supplement the MCBS sample. However, Hispanics continue to be oversampled with respect to non-Hispanics in the MCBS sample as in past years. A gradual increase in the Hispanic portion of the MCBS sample is planned over the next several years, with the goal of fully replacing the previously sampled Puerto Rican Hispanic beneficiaries.

#### 3.2 Questionnaires

**Questionnaire content changes:** There were a number of questionnaire sections that were revised in 2018. 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/Research-Statistics-Data-and-Systems/Research/MCBS/Questionnaires>.

##### 3.2.1 General

#### Community Questionnaire

The MCBS introduced several Community questionnaire updates in 2018 to enhance survey content and data quality, improve interviewer and respondent experience, and reduce respondent burden.

These changes included the addition of a new section measuring beneficiary's experience with chronic pain and the implementation of several new items, including those on weight management counseling, depression screening, smoking cessation, and how beneficiaries with Limited English Proficiency (LEP) usually communicate with health care providers. Additional details about questionnaire content and changes made in 2018 can be found in Section 3.2.2.

#### Facility Instrument

In 2018, one general maintenance questionnaire update was made to the Facility Instrument. In a number of sections, question wording and response options were updated to use the term "Intermediate Care Facility for Individuals with Intellectual Disabilities (ICF/IID)" instead of the outdated term "Intermediate Care Facility for the Mentally Retarded (ICF-MR)."

<sup>2</sup> Heretofore referred to as U.S. Hispanics.

### 3.2.2 Section-Specific Changes

#### Community Questionnaire

The MCBS introduced several changes to the Community questionnaire in 2018<sup>3</sup> to enhance survey content and data quality, improve interviewer and respondent experience, and reduce respondent burden.

#### Summary of Item- and Section-Level Questionnaire Revisions

##### Chronic Pain (CPQ)

- The Chronic Pain (CPQ) section was administered for the first time in Summer 2019 and is released in the 2018 LDS. Questionnaire items were developed by the National Pain Strategy (NPS) Population Research Working Group for inclusion in federal surveys. The CPQ begins with one item, 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.

##### Drug Coverage (RXQ)

- In Summer 2018, two questions that ask about the drug coverage gap, or "donut hole" (SC8I-DHPLAN and SC8L-DHTHISYR), were removed from the questionnaire because the information is available in administrative data. These two deleted items served as filter questions to determine when follow-up items about the drug coverage gap were applicable to the beneficiary. With the removal of the filter questions, all beneficiaries were asked follow-up items SC8M-DHSTART and SC8N-DHEND about the donut hole. However, some of these beneficiaries should have filtered out of these questions because they had not yet reached the start of the coverage gap or had private prescription drug coverage. In Summer 2019, the two questions about the coverage gap (SC8I-DHPLAN and SC8L-DHTHISYR) were added back to this series in their original form. As the questionnaire is administered the summer following the year of interest, the data collected in Summer 2019 are released in the 2018 LDS.

##### Health Status and Functioning (HFQ)

- In Fall 2018, two new screening items were added to the Health Status and Functioning (HFQ) questionnaire. Due to the longitudinal nature of the MCBS, the survey contains two variables for each screening item. Respondents participating in their Baseline interview are asked if the beneficiary has ever had the screening; a second similar version is administered to respondents participating in their Continuing interview and asks if the beneficiary has had the screening in the past year. The first new item was adapted from the Consumer Assessment of Healthcare Providers and Systems (CAHPS) for Accountable Care Organizations (ACOs) to collect information about depression screening (HFJ47-BASKDEPRS and HFJ48-CASKDEPRES). The other item was sourced from the National Health and Nutrition Examination Survey (NHANES) and asks about weight management counseling (HFJ45-BLOSWGHT and HFJ46-CLOSWGHT).

##### Nicotine and Alcohol Use (NAQ)

- One new item on tobacco cessation sourced from the National Health Interview Survey (NHIS) was added to the Nicotine and Alcohol Use (NAQ) section in Fall 2018 (NA25-BDOCSMOK or NA26-CDOCSMOK). For beneficiaries who report smoking every day or some days, this item asks whether a doctor or other health professional has ever talked to them about their smoking.

<sup>3</sup> 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/Research-Statistics-Data-and-Systems/Research/MCBS/Questionnaires>.

### Preventive Care (PVQ)

- For data collected in Summer 2019, routing logic in the Preventive Care (PVQ) section was updated to ask all respondents who did not report a flu shot in the Winter round (PVF1-FLUSHOT) the reason for not receiving the flu shot (PV2-FLUCODE).

### Usual Source of Care (USQ)

- Items PP1-REMINDAPPT and PP2-PREPARE in the USQ section, which ask about communication from doctor's offices prior to an upcoming appointment, were updated to add a new response option of "(996) NOT APPLICABLE / R DID NOT HAVE AN APPOINTMENT" to each of these two items.

## Facility Instrument

There were several changes to the Facility Instrument in 2018, including updates to question text, response options, programming logic, and text fills.

### Facility Questionnaire (FQ)

- In Fall 2018, the routing of billing rate questions was updated to maximize the amount of data collected. A Facility respondent who does not know the high billing rate at item FR3-HIGHRATE now has the opportunity to report a low billing rate at FR4-LOWRATE.

### Health Status (HS)

- Several changes were made to the Health Status (HS) section to better align it with the Long-Term Care Minimum Data Set (MDS). In Fall 2018, new programming logic was added for the Brief Interview for Mental Status (BIMS) assessment items to be consistent with how the MDS form handles incomplete BIMS assessments. Like the MDS, the Facility Instrument now collects information about the Staff Assessment for Mental Status in instances when the BIMS assessment was not completed. In Fall 2018, "Don't Know" and "Refused" response options were added to items HA29B-HA29BCOD, HA37AB-HA37ABCO, and HA37BB-HA37BBCO to ensure more accurate data entry in situations where no response options were selected on the MDS form.

## 3.3 Data Collection

**Rounds of Data Collection:** Beginning in 2018, the MCBS design was modified to eliminate the 12<sup>th</sup> interview; the 11<sup>th</sup> interview, conducted each winter round, will be the final (exit) interview. Therefore, the final interview for the 2014 Panel was conducted in Winter 2018 (Round 80).

## 3.4 Documentation

The 2018 *Data User's Guides* and *Methodology Report* were enhanced with new charts, tables, and descriptions. The enhanced content describes in further detail differences in data collection and data processing for beneficiaries living in facilities and in the community, explains differences between the Survey File and Cost Supplement data sources and population, and expands information on imputation and guidance for conducting analyses with community and facility data.

## 3.5 Data Processing

### New and revised content:

For the 2018 Survey File LDS, the MCBS implemented the following changes to segments:

- Two new segments are released: Chronic Pain (CHRNPAIN) and the Four-Year Survey File Longitudinal weights (LNG4WGTS).
- All Topical sections have updated weights for the 2018 Survey File ever enrolled population. See more information in Section 9.4: Weighting.
- Processing of the Area Deprivation Index (ADI) variables on the DEMO segment has been updated to use 2015 American Community Survey (ACS) datasets rather than 2000 Census-based ADI datasets.

The 2018 questionnaire changes also resulted in the following variables added to the annual releases:

### Exhibit 3.5.1: 2018 MCBS Content Additions

Location	Questionnaire Section	Variable	Description
<b>CHRNPAIN</b>	CPQ	MANGCBT	SP MANAGE PAIN- TALK THERAPY
<b>CHRNPAIN</b>	CPQ	MANGGROP	SP MANAGE PAIN- PEER SUPPORT GROUP
<b>CHRNPAIN</b>	CPQ	MANGMASG	SP MANAGE PAIN- MASSAGE
<b>CHRNPAIN</b>	CPQ	MANGMEDT	SP MANAGE PAIN- MEDITATION, GUIDED IMAGERY, OR OTHER RELAXATION TECHNIQUES
<b>CHRNPAIN</b>	CPQ	MANGPHYS	SP MANAGE PAIN- PT, REHAB, OT
<b>CHRNPAIN</b>	CPQ	MANGPROG	SP MANAGE PAIN- SELF-MANAGEMENT PROGRAM
<b>CHRNPAIN</b>	CPQ	MANGSPNE	SP MANAGE PAIN- SPINAL MANIP OR CHIROPRACTIC CARE
<b>CHRNPAIN</b>	CPQ	MANGYOGA	SP MANAGE PAIN- YOGA, TAI CHI, QI GONG
<b>CHRNPAIN</b>	CPQ	PAINABDM	HOW MUCH SP BOTHERED BY PAIN IN ABDOMEN, PELVIS, GENITALS
<b>CHRNPAIN</b>	CPQ	PAINAMNT	HOW MUCH PAIN SP HAD
<b>CHRNPAIN</b>	CPQ	PAINARMS	HOW MUCH SP BOTHERED BY PAIN IN HANDS, ARMS, SHOULDERS
<b>CHRNPAIN</b>	CPQ	PAINBACK	HOW MUCH SP BOTHERED BY PAIN IN BACK
<b>CHRNPAIN</b>	CPQ	PAINFAM	HOW OFTEN SP'S PAIN AFFECTED FAMILY AND SIGNIFICANT OTHERS
<b>CHRNPAIN</b>	CPQ	PAINHEAD	HOW MUCH SP BOTHERED BY HEADACHE, MIGRAINE, FACIAL PAIN
<b>CHRNPAIN</b>	CPQ	PAINLEGS	HOW MUCH SP BOTHERED BY PAIN IN HIPS, KNEES, FEET
<b>CHRNPAIN</b>	CPQ	PAINLIMT	HOW OFTEN SP'S PAIN LIMITED LIFE OR WORK ACTIVITIES
<b>CHRNPAIN</b>	CPQ	PAINMANG	WHAT EXTENT DID SP MANAGE PAIN TO DO THINGS THEY ENJOY
<b>CHRNPAIN</b>	CPQ	PAINOFTN	HOW OFTEN SP HAD PAIN
<b>CHRNPAIN</b>	CPQ	PAINTOTH	HOW MUCH SP BOTHERED BY TOOTHACHE OR JAW PAIN

Location	Questionnaire Section	Variable	Description
DEMO	CCDC, 2013 American Community Survey (ACS) data	ADINATNL	NATIONAL PERCENTILE ADI RANKING BASED ON CENSUS BLOCK GROUP
DEMO	CCDC, 2013 American Community Survey (ACS) data	ADISTATE	STATE DECILE ADI RANKING BASED ON CENSUS BLOCK GROUP
ACCSSMED	DVH	HVNDADVI	SP COULDN'T GET HEARING - ANOTHER DR REC NOT DOING IT
ACCSSMED	DVH	HVNDAFRD	SP COULDN'T GET HEARING - AFRAID
ACCSSMED	DVH	HVNDBUSY	SP COULDN'T GET HEARING - TOO BUSY
ACCSSMED	DVH	HVND COST	SP COULDN'T GET HEARING - COST
ACCSSMED	DVH	HVNDDIST	SP COULDN'T GET HEARING - OFFICE TOO FAR
ACCSSMED	DVH	HVNDFINE	SP COULDN'T GET HEARING - NOTHING WRONG
ACCSSMED	DVH	HVNDHOUR	SP COULDN'T GET HEARING - OFFICE NOT OPEN AT GOOD TIMES
ACCSSMED	DVH	HVNDINS	SP COULDN'T GET HEARING - INS DID NOT COVER PROCS
ACCSSMED	DVH	HVNDMONY	SP COULDN'T GET HEARING - DIDN'T WANT TO SPEND MONEY
ACCSSMED	DVH	HVNDOTHR	SP COULDN'T GET HEARING - OTHER SPECIFY
ACCSSMED	DVH	HVNDWORK	SP COULDN'T GET HEARING - COULDN'T TAKE OFF WORK
ACCSSMED	DVH	HVNEED	TIME SP COULDN'T GET HEARING
ACCSSMED	DVH	VUNDADVI	SP COULDN'T GET VISION - ANOTHER DR REC NOT DOING IT
ACCSSMED	DVH	VUNDAFRD	SP COULDN'T GET VISION - AFRAID
ACCSSMED	DVH	VUNDBUSY	SP COULDN'T GET VISION - TOO BUSY
ACCSSMED	DVH	VUNDCOST	SP COULDN'T GET VISION - COST
ACCSSMED	DVH	VUNDDIST	SP COULDN'T GET VISION - OFFICE TOO FAR
ACCSSMED	DVH	VUNDFINE	SP COULDN'T GET VISION - NOTHING WRONG
ACCSSMED	DVH	VUNDHOUR	SP COULDN'T GET VISION - OFFICE NOT OPEN AT GOOD TIMES
ACCSSMED	DVH	VUNDINS	SP COULDN'T GET VISION - INS DID NOT COVER PROCS
ACCSSMED	DVH	VUNDMONY	SP COULDN'T GET VISION - DIDN'T WANT TO SPEND MONEY
ACCSSMED	DVH	VUNDOTHR	SP COULDN'T GET VISION - OTHER SPECIFY
ACCSSMED	DVH	VUNDWORK	SP COULDN'T GET VISION - COULDN'T TAKE OFF WORK
ACCSSMED	DVH	VUNEEED	TIME SP COULDN'T GET VISION

Location	Questionnaire Section	Variable	Description
CHRNCOND	HFAQ	BASKDPRS	EVER HAD DR/HP ASK IF PERIOD OF TIME SP FELT SAD/EMPTY/DEPRESSED
CHRNCOND	HFAQ	BLOSWGHT	EVER BEEN TOLD BY DR/HP TO CONTROL/LOSE WEIGHT
NICOALCO	NAQ	DOCSMOK	EVER HAD DR/HP TALK TO SP ABOUT SMOKING
ADMNUTLS	N/A	H_ACTBPT	ACUTE INPATIENT BENEFICIARY PAYMENTS
ADMNUTLS	N/A	H_OIPBPT	OTHER INPATIENT BENEFICIARY PAYMENTS
ADMNUTLS	N/A	H_SNFBPT	SKILLED NURS FACILTY BENEFICIARY PAYMNTS
ADMNUTLS	N/A	H_HOPBPT	HOSPITAL OUTPATIENT BENEFICIARY PAYMENTS
ADMNUTLS	N/A	H_BPTDRG	PART B DRUG BENEFICIARY PAYMENTS
ADMNUTLS	N/A	H_EMBPT	EVALUATION & MGMNT BENEFICIARY PAYMENTS
ADMNUTLS	N/A	H_PHYBPT	PART B PHYSICIAN BENEFICIARY PAYMENTS
ADMNUTLS	N/A	H_ORPBPT	OTHER PROCEDURES BENEFICIARY PAYMENTS
ADMNUTLS	N/A	H_DMEBPT	DURABLE MEDICAL EQUIPMENT BENE PAYMENTS
ADMNUTLS	N/A	H_BPTOTH	OTHER PART B CARRIER BENEFICIARY PAYMNTS
ADMNUTLS	N/A	H_PTDBPT	PART D BENEFICIARY PAYMENTS
ADMNUTLS	N/A	H_ASCBPT	AMBULATORY SURGERY CENTER BENE PAYMENTS
ADMNUTLS	N/A	H_ANEBPT	ANESTHESIA BENEFICIARY PAYMENTS
ADMNUTLS	N/A	H_DIABPT	DIALYSIS BENEFICIARY PAYMENTS
ADMNUTLS	N/A	H_IMGBPT	IMAGING BENEFICIARY PAYMENTS
ADMNUTLS	N/A	H_TSTBPT	TESTS BENEFICIARY PAYMENTS
OASIS	N/A	HHASMTID	HHA ASSESSMENT INTERNAL ID
RESTMLN	N/A	D_BEGr	SITUATION BEGIN DATE
RESTMLN	N/A	D_ENDr	SITUATION END DATE
RESTMLN	N/A	D_NUMEVT	# OF FFS EVENTS
RESTMLN	N/A	D_NUMRES	# OF RESIDENTIAL SETTINGS
RESTMLN	N/A	D_TCDAYS	# OF COMMUNITY DAYS
RESTMLN	N/A	D_TFDAYS	# OF FACILITY DAYS
INTERV	N/A	INTVFLG	INTERVIEW TIMEFRAME
HHCHAR	N/A	SPMOVED	SP ADDRESS CHANGE DUE TO MOVE

**Data editing and imputation procedures:** MCBS data files undergo thorough editing and quality control checks prior to release. For more detailed information regarding data editing and imputation procedures conducted for the 2018 LDS releases, please consult the forthcoming *2018 Data User's Guide: Cost Supplement File* and *2018 MCBS Methodology Report* available on the CMS MCBS website.

### Weighting:

The 2018 Survey File LDS includes a two-year longitudinal weight for analyses using 2017 and 2018 data, a three-year longitudinal weight for analyses using 2016 and 2018 data, and a four-year longitudinal weight for

analyses using 2015 and 2018 data. The four-year longitudinal weights are new, having been reintroduced for 2018.

In previous years, supplied weights for the Topical data corresponded to two MCBS populations: the Survey File continuously enrolled population and the Cost Supplement ever enrolled population. This year, new Topical weights corresponding to the Survey File ever enrolled population are also available.

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

**Imputation:**

For 2018, the Income and Assets imputation made greater use of survey reported income from earlier years, reducing the reliance on hot deck imputation based solely on demographic variables.

## 4. SURVEY OVERVIEW

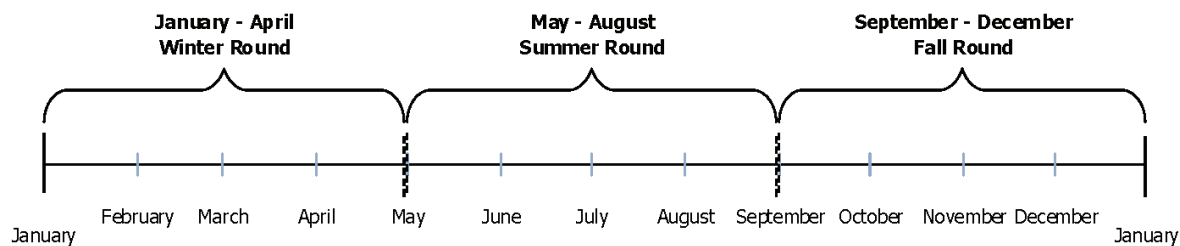
### 4.1 Design of MCBS

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 2018 MCBS data releases reflect data collected from January 2018 through early January 2019, as well as Topical sections, income and assets data, and chronic pain information collected through the Summer 2019.<sup>4</sup> Exhibit 4.1.1 depicts an MCBS data collection year and the typical span of the rounds.

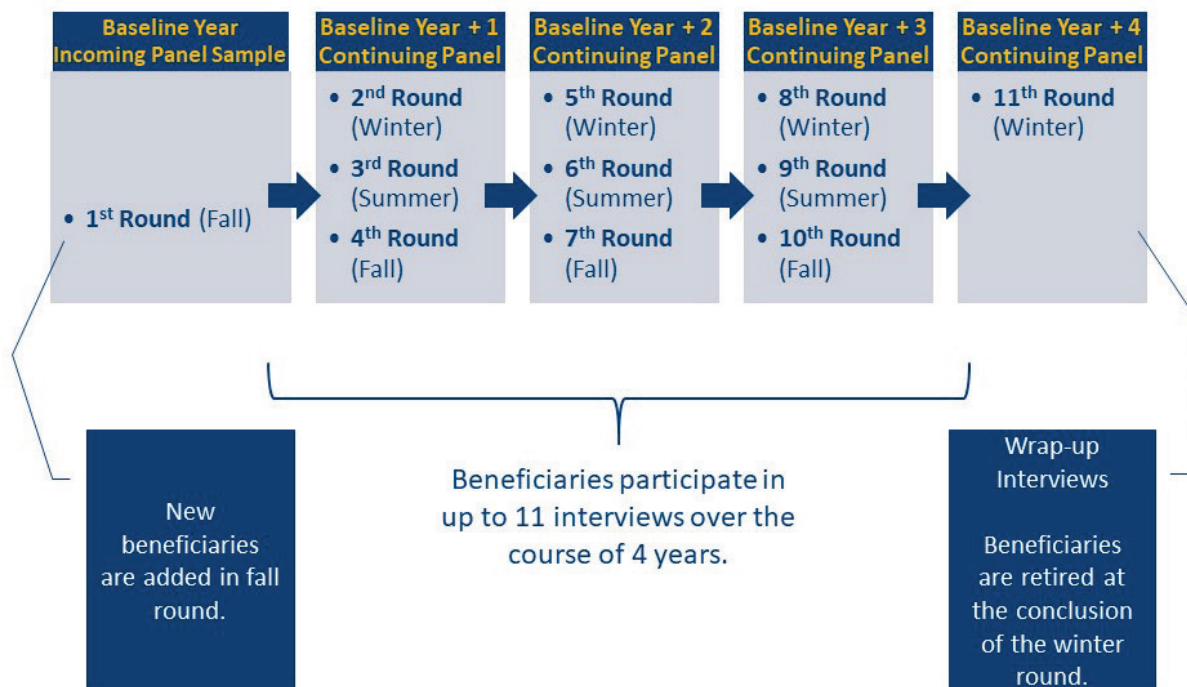
**Exhibit 4.1.1:** Typical MCBS Data Collection Year



Initial interviews of newly-selected respondents take place in the fall round. Often the fall round begins early (i.e., 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. For example, the fall round usually extends into early January to allow for the completion of interviews that may have been postponed due to the holiday period.

Subsequent rounds, which occur every four months, involve re-interviewing of the same respondent (or appropriate proxy respondents) until they have completed four years of participation (up to 11 interviews in total). Interviews are conducted regardless of whether the respondent 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 respondents selected to be in the MCBS sample and Appendix B provides a list of all rounds by data collection year.

<sup>4</sup> 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.

**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.<sup>5</sup> Each MCBS panel, an annual statistical sample of all Medicare enrollees, is interviewed up to three times per year over four consecutive years creating a continuous profile of selected beneficiaries' health care experiences.<sup>6</sup> 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 respondents across all panels participating in the survey annually. Please see Section 6: Sampling for more information on the sample design selection.

<sup>5</sup> 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.

<sup>6</sup> 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:** 2014-2018 MCBS Rotating Panel Design

Data Collection Schedule			Panel				
Data Year	Season	Round#	2014	2015	2016	2017	2018
2014	Winter	68					
	Summer	69					
	Fall	70					
2015	Winter/Summer*	71/72					
	Fall	73					
2016	Winter	74					
	Summer	75					
	Fall	76					
2017	Winter	77					
	Summer	78					
	Fall	79					
2018	Winter	80					
	Summer	81					
	Fall	82					

\*The Summer and Winter Rounds in 2015 were combined due to a contract transition.

### 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 respondents completing a final interview and exiting the MCBS in the prior winter round. Respondents new to the MCBS and introduced 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 90 percent of the interviews take place in the respondent or proxy's own residence or in a neutral interview location, such as a library or public venue. These interviews are called Community interviews; the remaining 10 percent of the interviews are for beneficiaries living in a facility. Over the course of a four-year period, it is not uncommon for respondents 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 respondent directly; instead, specially trained Facility interviewers administer the survey with Facility administrative staff.

## 4.4 Interviewing and Training Procedures

### 4.4.1 Overview of Data Collection

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

### Overview of recruitment of beneficiaries and scheduling procedures

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

### Computer-Assisted Personal Interviewing (CAPI)

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

### 4.4.2 Interviewer Training

Nationally, the MCBS employs an average of approximately 200<sup>7</sup> field interviewers, who participate in a combination of several targeted training initiatives and careful coaching and monitoring activities throughout data collection. Each training is customized to the level of experience of the interviewer (new to MCBS or MCBS-experienced), the type of interview (Community or Facility), the type of sample (Incoming Panel or Continuing), and the unique requirements of each round (changing questionnaire sections or data collection protocols). Field interviewers who are new to MCBS are always trained in person; experienced field interviewers participate in a periodic in-person training program and receive continuous online refresher training. Weekly field memos issued to all field managers and field interviewers cover important data collection tips, provide answers to interviewer questions, and offer reminders about how to handle complicated scenarios.

### 4.4.3 Privacy and Data Security

Field interviewer training stresses the importance of maintaining respondent privacy, and project protocols are documented within the Field Interviewer manual. Field outreach and contacting procedures also maintain and ensure confidentiality. These procedures include the utilization of standard computer security protocol (dual

<sup>7</sup> The fall round starts with a target of 230 field interviewers which, over the course of the year, is reduced due to staff turnover. Each summer, a cohort of new interviewers is hired for the MCBS.

authentication password protection for each interviewer laptop) and restrictions on submitting personally identifiable information (PII) through electronic mail. All MCBS survey staff directly involved in data collection and/or analysis activities are required to sign a Non-Disclosure Agreement and a confidentiality agreement.

NORC and CMS are committed to protecting respondent confidentiality and privacy, and both organizations diligently uphold provisions established under the Privacy Act of 1974, the NORC Institutional Review Board (IRB), the Office of Management and Budget (OMB), and the Federal Information Security Management Act of 2002. As stated in the MCBS OMB documentation, the information collected for MCBS is protected by NORC and by CMS. Respondent data are used only for research and statistical purposes. As required under the Privacy Act of 1974, identifiable information is not disclosed or released without the consent of the individual or the establishment, except to those involved in research (Public Law 93-579).

## 4.5 Completed Interviews

Exhibit 4.5 lists the number of completed interviews for the Fall 2018 Continuing (2015, 2016, and 2017) and Incoming (2018) Panels by age strata. Under the rotating panel design, the beneficiaries selected in Fall 2014 exited the study at the conclusion of the Winter 2018 round.

**Exhibit 4.5:** 2018 MCBS Fall Round Completed Interviews: Continuing and Incoming Panels

Age Category as of 12/31/2018	2015 Panel	2016 Panel	2017 Panel	2018 Panel	Total
<b>Under 45 years</b>	123	248	295	650	1,316
<b>45-64 years</b>	149	236	297	491	1,173
<b>65-69 years</b>	138	370	601	1,158	2,267
<b>70-74 years</b>	282	479	587	869	2,217
<b>75-79 years</b>	321	439	658	952	2,370
<b>80-84 years</b>	329	461	614	990	2,394
<b>85+ years</b>	426	696	691	1,046	2,859
<b>Total</b>	1,768	2,929	3,743	6,156	14,596

SOURCE: 2018 MCBS Internal Sample Control File

## 4.6 Item Non-Response

As in any other survey, some respondents could not, or would not, supply answers to some questions.<sup>8</sup> Item non-response rates are generally low in the MCBS data, but the analyst still needs to be aware of the missing data and be cautious about patterns of non-response.<sup>9</sup> Some of the missing data are attributable to the fact that some of the Community interviews and all of the Facility interviews are conducted through a proxy respondent. In other words, the respondent may not have had knowledge of the information sought on the sample person. In other situations, the respondent may have simply refused to answer.

<sup>8</sup> This is different from when an individual refuses to participate in the survey altogether, which is called unit non-response. Unit non-response is discussed in detail in the *MCBS Methodology Report*, Section 9.

<sup>9</sup> In the LDS files, item non-response types are indicated by missing type codes in SAS, including refusal to answer, don't know the answer, and invalid skip. The code .D represents a "don't know" response, the code .R represents a "refused" response, and the code .N represents an "invalid skip" response.

## 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, 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.

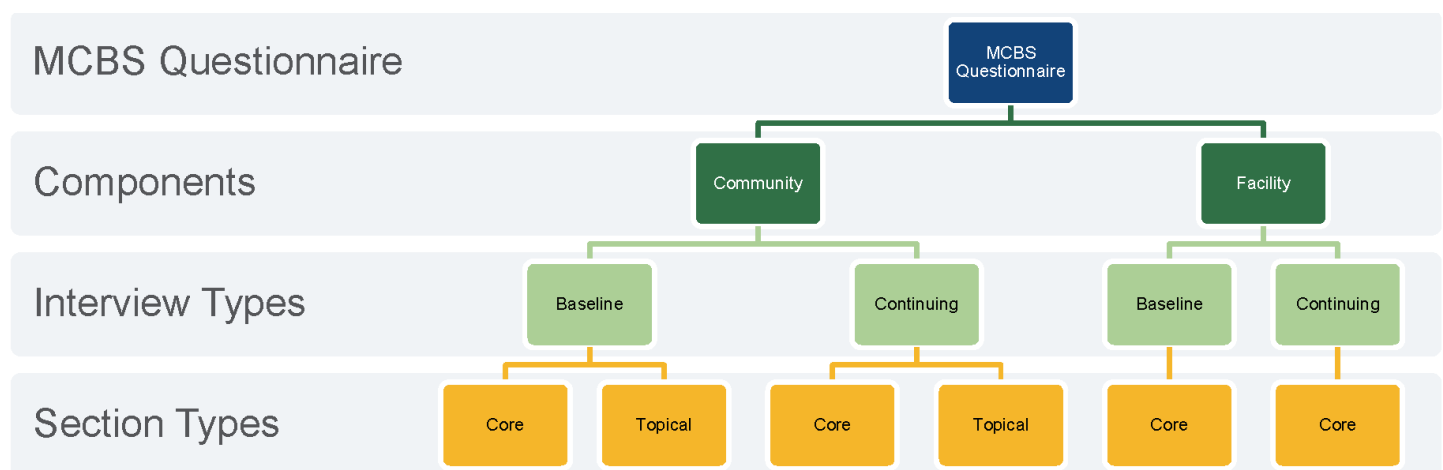
- **Community Component:** Survey of beneficiaries living in the community at the time of the interview (i.e., their residence or a household). An interview may be conducted with the beneficiary or a proxy.
- **Facility Component:** Survey of respondents living in facilities, such as long-term care nursing homes or other institutions at the time of the 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 to beneficiaries new to the study; administered in the fall of the year they are selected into the sample (interview #1).
- **Continuing:** The questionnaire administered to beneficiaries as they 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 2018 Core and Topical sections.

**Exhibit 5.1:** MCBS Questionnaire Overview



## 5.2 Community Questionnaire

The content of the MCBS Community Questionnaire consists of Core and Topical sections. Core sections include the standard opening and closing sections covering interview characteristics and socio-demographics, health insurance, utilization, cost, experiences with care, and health status sections. The questionnaire sections in each of these categories may be asked each round or seasonally (fall, winter, summer). Topical sections in the Community Questionnaire include information about housing characteristics, health behaviors, and knowledge and decision-making.

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

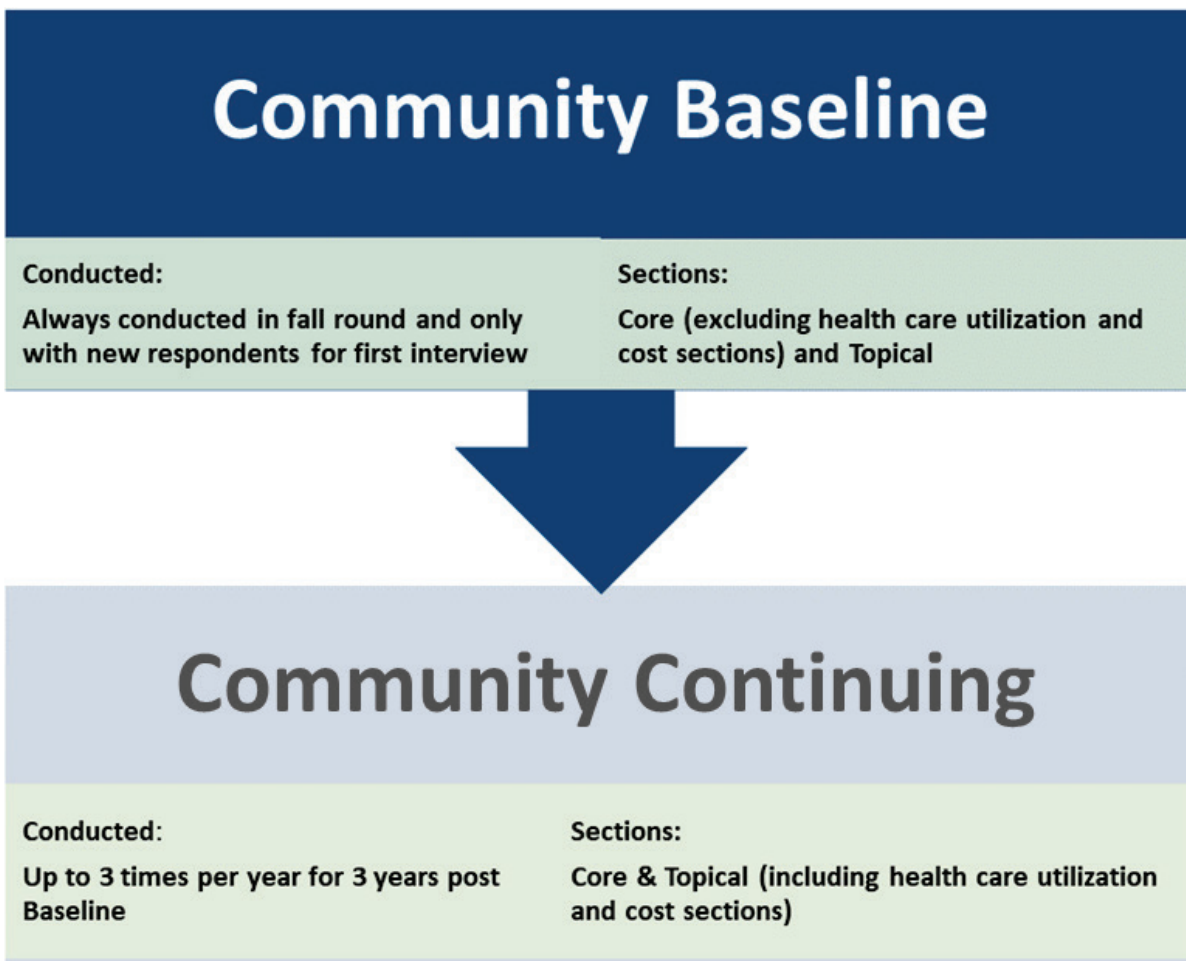
The first Community interview conducted with Incoming Panel respondents is referred to as the Baseline interview. This interview is always conducted in the fall round and consists of a combination of Core and Topical sections. However, this first interview does not include Core sections that collect health care utilization and cost data. The respondent's 2<sup>nd</sup> through 11<sup>th</sup> 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 essentially 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):**

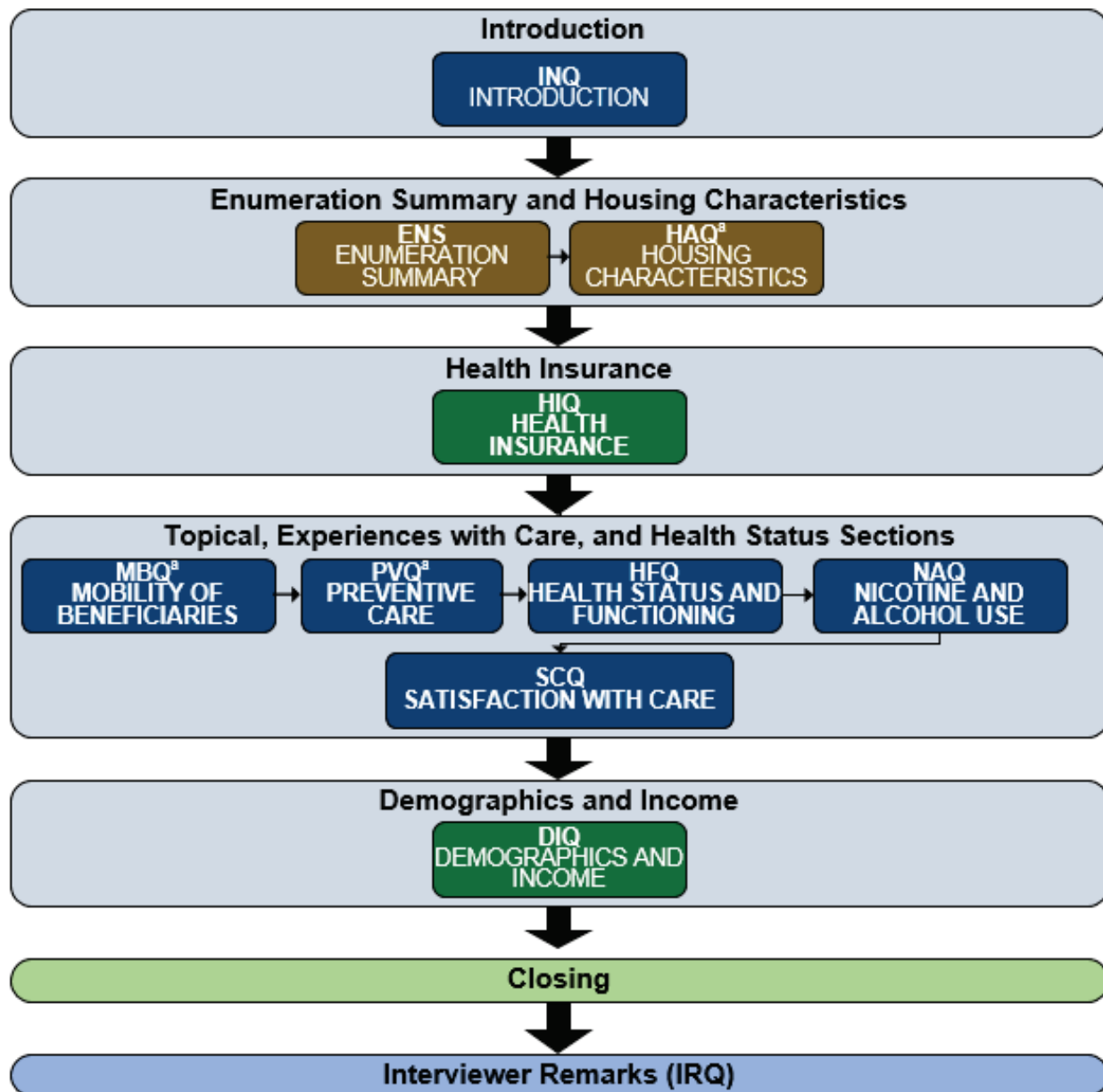
- **Community Baseline questionnaire**
- **Community Continuing questionnaire**

The data collection year includes three rounds of survey administration that occur within a calendar year. That is, for the 2018 data collection year, the rounds are Winter, Summer, and Fall administered from January 2018 through early January 2019. In contrast, the data year includes data collected in the calendar year as well as any data from prior or future calendar years that are included in the LDS's. That is, the data year includes the calendar year and data collected over three years. Some data collected in the previous year are pulled forward to fill in data for the current data year. This happens 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 the following year because the reference period for the questionnaire extends back into the data year.

In this section, data users will note that exhibit titles will indicate either the data collection year, which refers to the three rounds of survey administration that occur within the calendar year, or the data year, which refers to the data collected over three years that are included in the LDS's.

**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 the same beneficiary over the course of their participation in the survey, so establishing a positive relationship is critical during this first interview. Exhibit 5.2.1 depicts the sections and flow of the Community Baseline interview.

**Exhibit 5.2.1:** 2018 Data Collection Year MCBS Community Questionnaire Flow for Baseline Interview

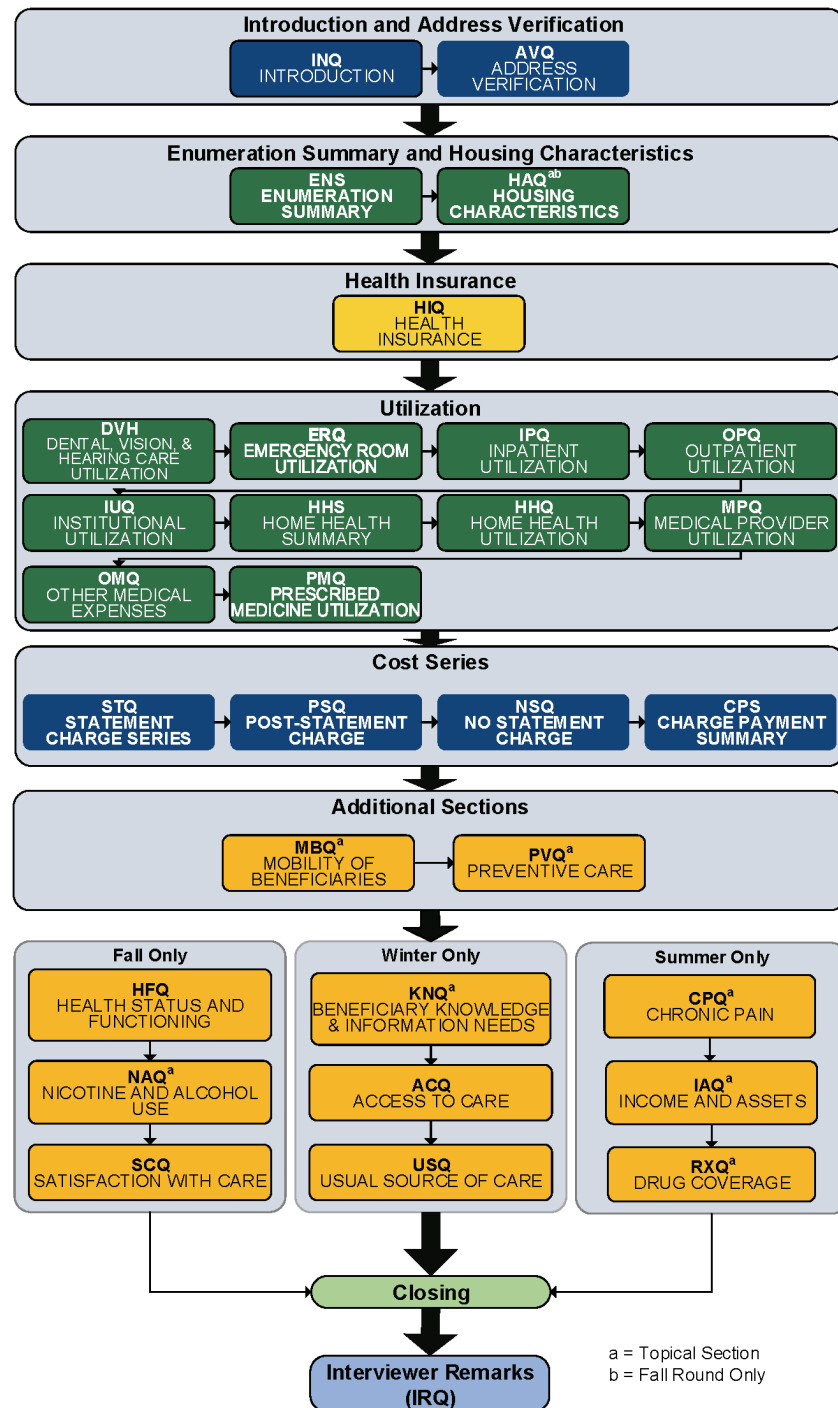
a = Topical Section

### 5.2.2 Continuing Interview

The Continuing interview consists of Core sections that focus on use of medical services and the resulting costs, and are asked in essentially the same way each and every time a section is 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 costs from the 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 that cover subjects such as mobility or drug coverage. Exhibit 5.2.2 depicts the sections and flow of the Community Continuing interview. All sections are considered "Core" sections unless otherwise noted.

**Exhibit 5.2.2:** 2018 Data Collection Year MCBS Community Questionnaire Flow for Continuing Interview



### 5.2.3 Core Questionnaire Sections

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 2018, only persons in the 2015, 2016, and 2017 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:** 2018 Data Year MCBS Community Core Sections by Data File and Data Collection Schedule\*

Section Group	Abbr.	Section Name	LDS <sup>§</sup>	Data Collection Schedule
<b>Socio-Demographics</b>	IAQ	Income and Assets	SF	Summer 2019**
	DIQ	Demographics/Income	SF	Fall 2018, Baseline Interview
<b>Health Insurance</b>	HIQ	Health Insurance	SF	All Seasons
<b>Utilization</b>	DUQ	Dental 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
	HHS	Home Health Summary <sup>±</sup>	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
	STQ	Statement Cost Series	CS	All Seasons
	PSQ	Post-Statement Charge	CS	All Seasons
<b>Cost</b>	NSQ	No Statement Charge	CS	All Seasons
	CPS	Charge Payment Summary <sup>±</sup>	CS	All Seasons
	ACQ	Access to Care	SF	Winter 2019**
	SCQ	Satisfaction with Care	SF	Fall 2018
<b>Experiences with Care</b>	USQ	Usual Source of Care	SF	Winter 2019**
	HFQ	Health Status and Functioning	SF	Fall 2018

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.

<sup>±</sup>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.

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

### 5.2.4 Topical Questionnaire Sections

Exhibit 5.2.4 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, and RXQ). Annually, special non-response adjustment weights are included within the segments for use in analysis when data are not collected within the same calendar year (see Exhibit 5.2.5).

**Exhibit 5.2.4:** 2018 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 2018
Social Determinants of Health or Health Behaviors	CPQ	Chronic Pain	SF	Summer 2019
	MBQ	Mobility of Beneficiaries	SF	Fall 2018, Winter 2019, and Summer 2019 <sup>±</sup>
	NAQ	Nicotine and Alcohol Use	SF	Fall 2018
	PVQ	Preventive Care	SF	Fall 2018, Winter 2019, and Summer 2019 <sup>±</sup>
	IAQ	Food Insecurity items	SF	Summer 2019 <sup>**±</sup>
Knowledge and Decision Making	KNQ	Beneficiary Knowledge and Information Needs	SF	Winter 2019 <sup>±</sup>
	RXQ	Drug Coverage	SF	Summer 2019 <sup>±</sup>

SOURCE: MCBS Community Questionnaire

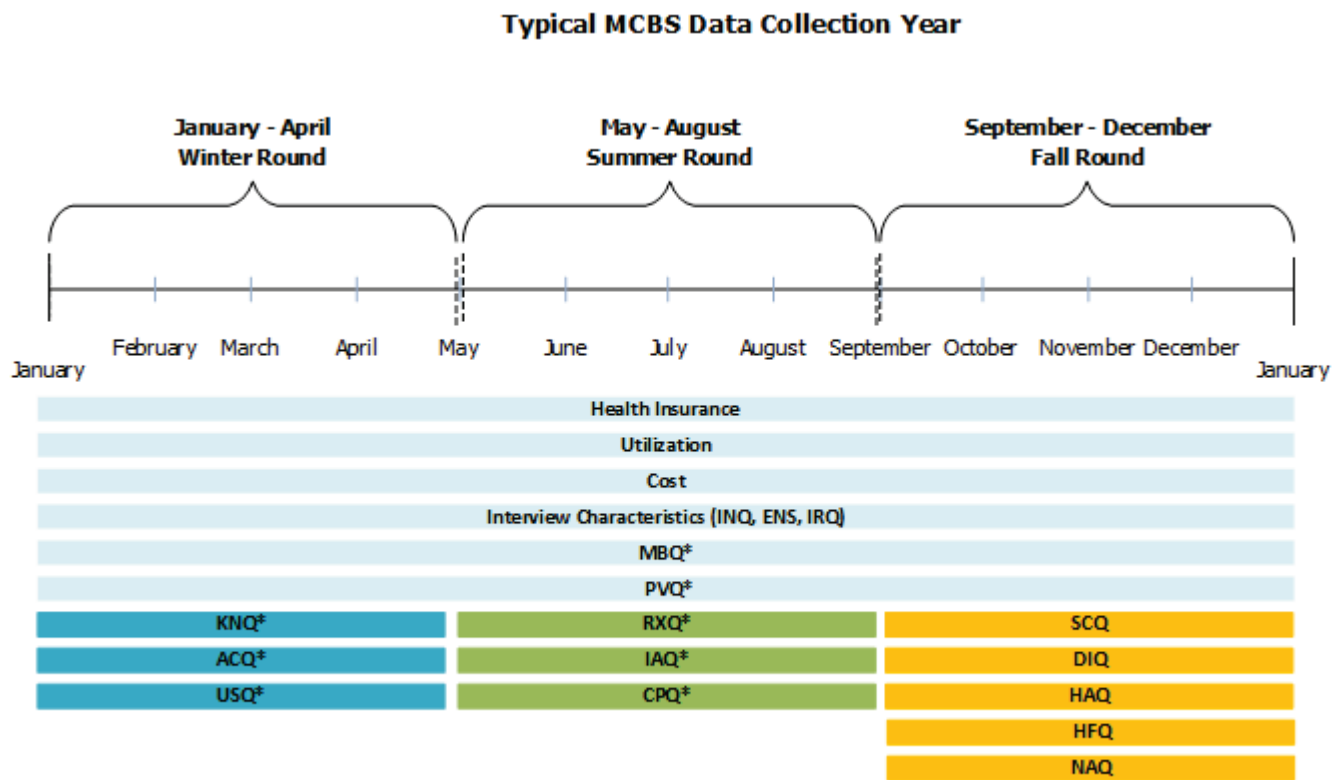
\*Limited Data Set (LDS) indicates the file where the questionnaire data appears (i.e., SF = Survey File, CS = Cost Supplement File).

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

<sup>±</sup>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.

### 5.2.5 Community Questionnaire Section Rotation within a Data Year

Exhibit 5.2.5 presents the MCBS Questionnaire section rotation schedule for 2018. Thus, the 2018 MCBS data releases reflect data collected from January 2018 through the first week in January 2019, as well as Topical sections collected through Summer 2019.

**Exhibit 5.2.5:** 2018 Data Collection Year MCBS Community Questionnaire Section Rotation

\*Fielded in 2019, but given the reference period is 2018, data are included in the 2018 LDS's.

### 5.3 Facility Instrument

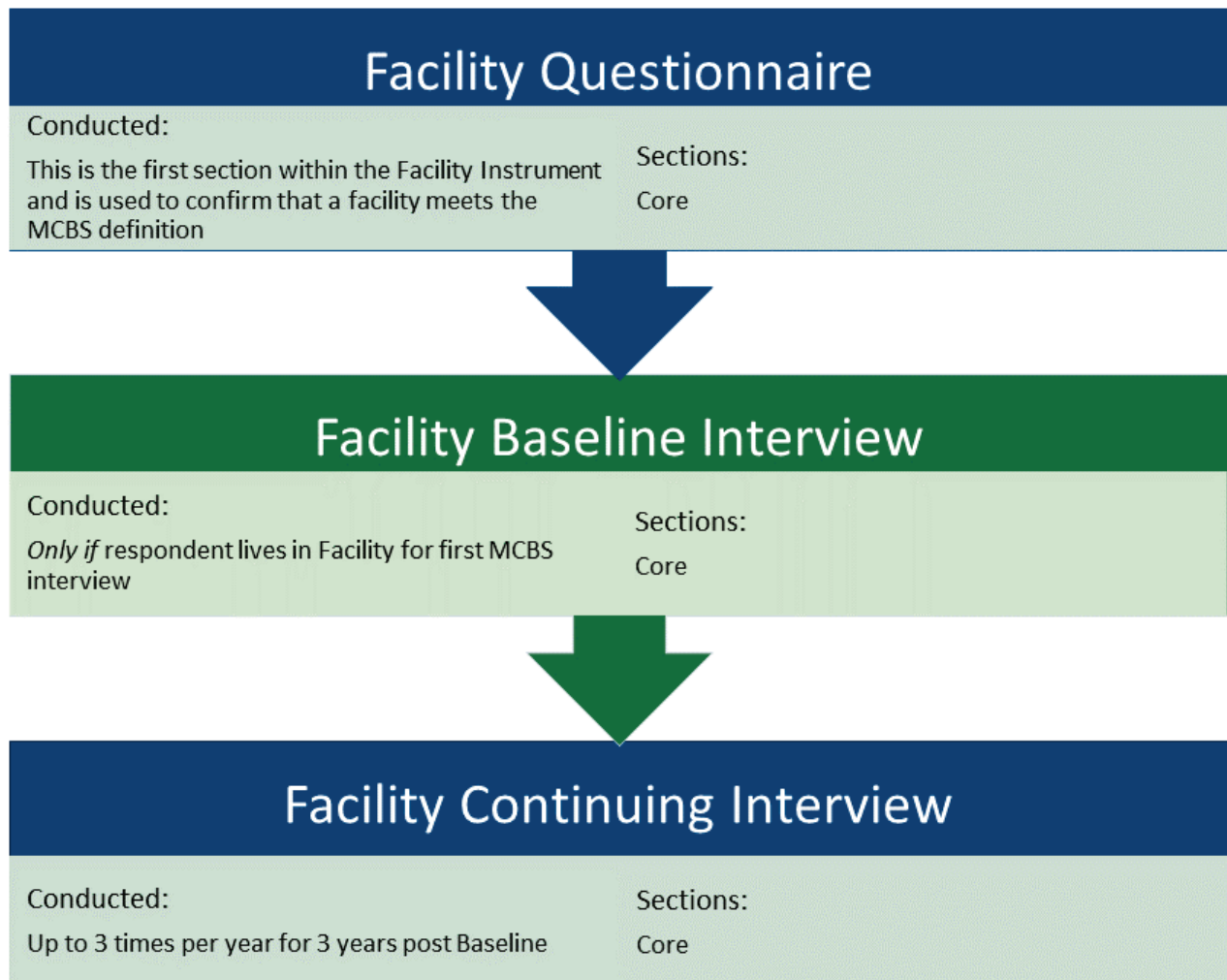
In addition to collecting information from respondents living in the community, the MCBS collects information at the institutional level if the beneficiary is living in a facility at the time of the interview. Information is obtained only by interviewing facility staff; the beneficiary is never interviewed 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 2<sup>nd</sup> through 11<sup>th</sup> 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.

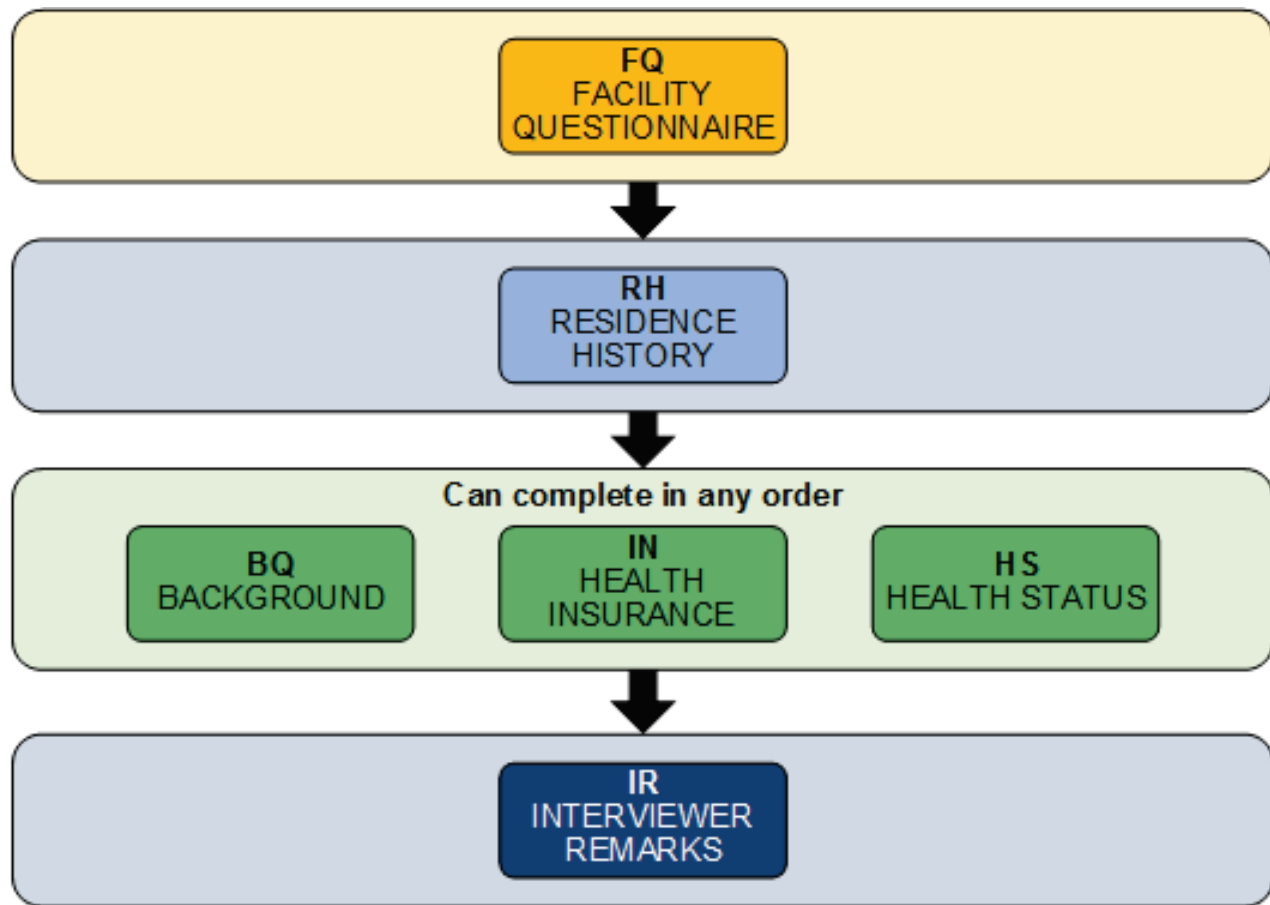
**The Facility Instrument consists of the following components (see Exhibit 5.3):**

- Facility Questionnaire
- Facility Baseline interview
- Facility Continuing interview

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.

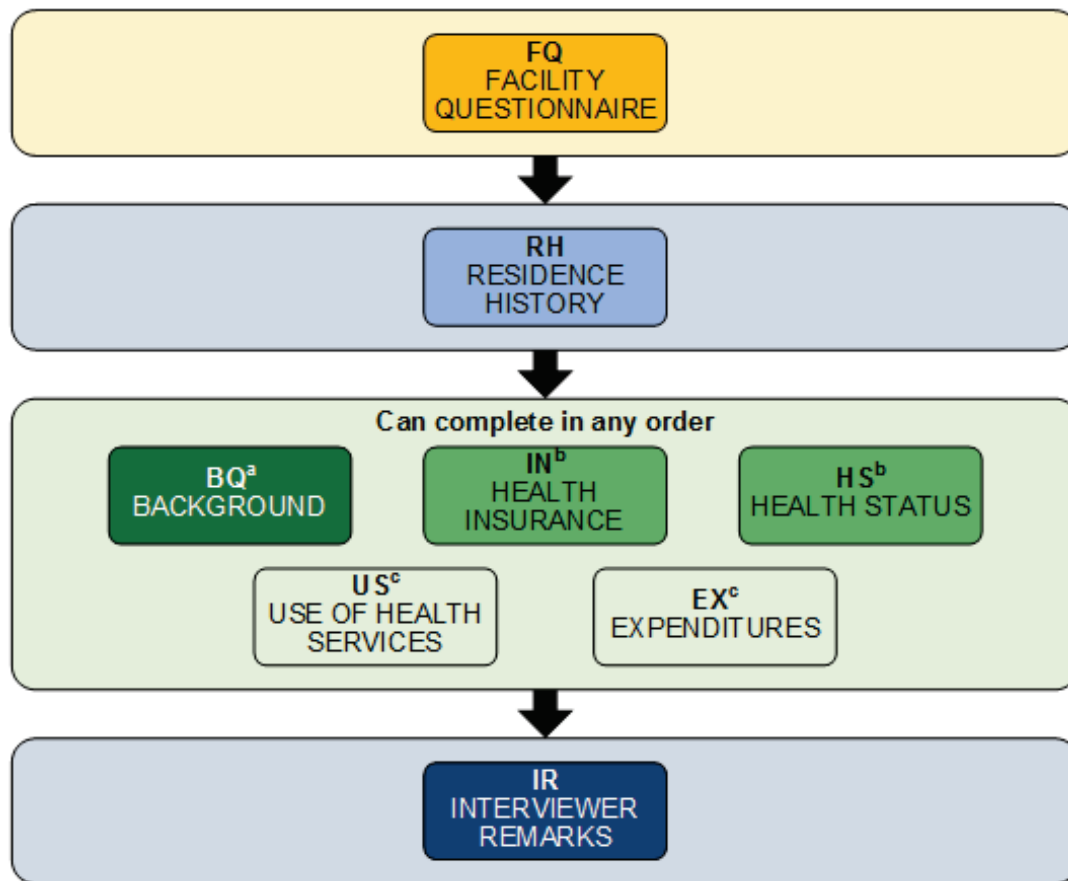
**Exhibit 5.3:** Overview of the MCBS Facility Instrument**5.3.1 Facility Baseline Interview**

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.

**Exhibit 5.3.1:** 2018 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. 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:** 2018 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 residing 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

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 2<sup>nd</sup> through the 11<sup>th</sup> rounds); however, beneficiaries new to a facility receive additional Core sections.

Similarly 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:** 2018 Data Year MCBS Facility Core Sections by Data File and Data Collection Schedule\*

Section Group	Abbrev	Section Name	LDS <sup>§</sup>	Data Collection Schedule
Facility Characteristics	FQ	Facility Questionnaire	SF	All seasons
Socio-Demographics	RH	Residence History	SF	All seasons
	BQ	Background	SF	Fall 2018, Baseline Interview**
Health Insurance	IN	Health Insurance	SF	Fall 2018 <sup>±</sup>
Utilization	US	Use of Health Services	CS	All seasons
	PM	Prescribed Medicines	CS	All seasons
Cost	EX	Expenditures	CS	All seasons
Health Status	HS	Health Status	SF	Fall 2018 <sup>±</sup>

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., Interview Remarks (IR)).

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

<sup>±</sup>The IN and HS sections are also administered to Community to Facility and Facility to Facility cases each season.

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

## 6. SAMPLING

### 6.1 Medicare Population Covered by the 2018 MCBS Data

The MCBS data releases are a reflection of enrolled Medicare beneficiaries residing in the continental United States.<sup>10</sup> 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 2018 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 2018. A subset of these beneficiaries represent a random cross-section of all beneficiaries who were continuously enrolled from January 1, 2018 up to and including interviews conducted during Fall 2018. 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 2018 as well as those who dis-enrolled or died prior to their fall interview.<sup>11</sup> The ever enrolled population includes beneficiaries who were enrolled in Medicare for at least one day at any point during 2018.
- The continuously enrolled population represents only those individuals continuously enrolled in Medicare from January 1, 2018 up to and including their fall interview; this specifically excludes beneficiaries who enrolled during the calendar year 2018 and those who dis-enrolled or died prior to their fall interview. The concept of continuously enrolled is consistent with the concept of being exposed or “at risk” for using services up to and including their fall interview.

The 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 2018 LDS files, the 2015, 2016, 2017 and 2018 panels). Exhibit 6.1 shows the composition of each of the four panels included in the 2018 data files.

**Exhibit 6.1:** 2018 MCBS Composition of Panels in LDS Data Files

Data Year (Fall)	Number of Beneficiaries Selected
2015	8,621
2016	12,145
2017	11,623
2018	11,523

<sup>10</sup> Prior to 2017, Puerto Rico was also included as part of the MCBS sampling geography. Beginning in 2017, Puerto Rico was removed, and only beneficiaries living in the continental U.S. were eligible to be sampled for the MCBS. Beginning in 2018, all data collection in Puerto Rico was discontinued. The decision to remove Puerto Rico from the sample was based largely on the fact that Medicare in Puerto Rico is very different and difficult to compare to (or combine analytically with) Medicare in the U.S.

<sup>11</sup> Note that data collection for beneficiaries who enrolled during 2018 and died in 2018 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. 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, 2018) for male and female beneficiaries.

**Exhibit 6.1.1:** 2018 Total Estimated Number of Medicare Beneficiaries by Sex and Race

Group	Subgroup	Continuously Enrolled	Ever Enrolled
<b>Overall Total</b>		55,830,875	60,927,980
<b>Sex</b>	Male Total	25,457,794	27,876,250
	Female Total	30,373,081	33,051,730
<b>Race</b>	White non-Hispanic Total	41,844,104	44,258,789
	Black non-Hispanic Total	5,729,201	5,915,932
	Hispanic Total	3,758,078	3,909,330
	Other Total*	4,499,492	6,843,929

SOURCE: 2018 Survey File and Sample Control File, weighted counts.

\*The 'Other' race category includes other races, more than one race, and unknown race.

**Exhibit 6.1.2:** 2018 Estimated Number of Male Medicare Beneficiaries by Race and Age

Race	Age as of 12/31/2018	Continuously Enrolled	Ever Enrolled
<b>White non-Hispanic</b>	0-44	511,006	529,497
	45-64	2,445,336	2,382,984
	65-69	3,962,441	4,528,571
	70-74	4,766,503	4,847,313
	75-79	3,270,347	3,361,660
	80-84	2,150,572	2,303,471
	85+	1,907,517	2,190,428
<b>Black non-Hispanic</b>	0-44	169,177	190,381
	45-64	473,401	445,450
	65-69	527,127	588,707
	70-74	551,115	546,298
	75-79	347,495	336,615
	80-84	194,613	198,247
	85+	132,373	159,271
<b>Hispanic</b>	0-44	90,031	92,719
	45-64	268,476	255,090
	65-69	365,078	395,939
	70-74	415,409	429,654
	75-79	266,078	264,121
	80-84	144,121	163,105
	85+	142,655	177,724
<b>Other*</b>	0-44	109,699	199,718
	45-64	206,210	442,902
	65-69	1,312,916	2,075,107
	70-74	447,956	469,167
	75-79	158,753	158,552
	80-84	68,454	74,498
	85+	52,935	69,062

SOURCE: 2018 Survey File and Sample Control File, weighted counts.

\*The 'Other' race category includes other races, more than one race, and unknown race.

**Exhibit 6.1.3:** 2018 Estimated Number of Female Medicare Beneficiaries by Race and Age

<b>Race</b>	<b>Age as of 12/31/2018</b>	<b>Continuously Enrolled</b>	<b>Ever Enrolled</b>
<b>White non-Hispanic</b>	0-44	410,111	427,060
	45-64	2,139,494	2,108,587
	65-69	4,866,393	5,474,032
	70-74	5,558,251	5,667,351
	75-79	4,197,321	4,184,291
	80-84	2,670,528	2,801,625
	85+	2,988,284	3,451,921
<b>Black non-Hispanic</b>	0-44	125,601	135,474
	45-64	680,700	635,376
	65-69	771,546	846,935
	70-74	595,807	593,919
	75-79	492,344	511,553
	80-84	300,876	315,249
	85+	367,026	412,457
<b>Hispanic</b>	0-44	62,862	66,835
	45-64	313,342	307,573
	65-69	390,229	404,045
	70-74	479,235	502,001
	75-79	352,759	346,890
	80-84	235,539	256,003
	85+	232,266	247,629
<b>Other*</b>	0-44	91,606	146,376
	45-64	185,563	347,479
	65-69	1,083,743	2,057,042
	70-74	434,622	426,339
	75-79	158,354	159,657
	80-84	86,614	89,341
	85+	102,068	128,688

SOURCE: 2018 Survey File and Sample Control File, weighted counts.

\*The 'Other' race category includes other races, more than one race, and unknown race.

## 6.2 Targeted Population and Sampling Strata

Historically, 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 January 1 of the applicable sample-selection year, and whose address on the Medicare files is in one of the 48 contiguous states (excludes Alaska and Hawaii), the District of Columbia, or Puerto Rico. Beginning in 2015, the targeted population for the MCBS consisted of Part A and/or Part B enrollees as of December 31 of the sample-selection year. For example, for Fall Rounds 2015, 2016, 2017, and 2018 (the four rounds in which the 2015, 2016, 2017, and 2018 Panels, included in the 2018 MCBS data, were selected), the targeted population included those individuals enrolled as of December 31 of 2015, 2016, 2017, and 2018, respectively. Beginning in 2017, Puerto Rico was removed from the MCBS sample; thus, the MCBS sample was selected entirely from the continental U.S. and the District of Columbia beginning with the 2017 Panel.

The universe of beneficiaries for the MCBS is divided into seven sampling strata based on age as of a specified date during the calendar year of the data release. Beginning in 2015, this date was moved from July 1 to 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. Beginning in 2015, the strata were expanded to separate U.S. Hispanic, U.S. non-Hispanic, and Puerto Rican beneficiaries by age group. Beginning in 2017, with the removal of Puerto Rico, only two Hispanic strata remained (Hispanic and non-Hispanic). The 14 strata in 2018 include those depicted in Exhibit 6.2.1.<sup>12</sup>

**Exhibit 6.2.1:** 2018 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 2015, 2016, 2017, and 2018 Panels, beneficiaries residing within the U.S. who were Hispanic (based on a Hispanic ethnicity classification code in the Medicare enrollment data; see Eicheldinger<sup>13</sup> for more details) were oversampled. Exhibit 6.2.2 displays the beneficiaries selected as part of the 2018 Panel, by age and ethnicity.

<sup>12</sup> Note that the MCBS surveys beneficiaries living in noninstitutionalized (e.g., households) and institutionalized (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.

<sup>13</sup> Eicheldinger, C. "More Accurate Racial and Ethnic Codes for Medicare Administrative Data," *Health care financing review* 29, no. 3.

**Exhibit 6.2.2:** 2018 Panel of Selected Beneficiaries by Hispanic and Non-Hispanic Ethnicity Classification and Age Category

Age Category as of 12/31/2018	TOTAL Sample Size	TOTAL Weighted	Hispanic Sample Size	Hispanic Weighted	Non-Hispanic Sample Size	Non-Hispanic Weighted
<b>Under 45 years</b>	1,177	1,815,033	115	164,458	1,062	1,650,575
<b>45-64 years</b>	860	6,973,662	120	552,538	740	6,421,124
<b>65-69 years</b>	2,216	15,874,432	253	788,538	1,963	15,085,894
<b>70-74 years</b>	1,606	13,645,245	181	1,067,761	1,425	12,577,483
<b>75-79 years</b>	1,745	9,771,581	198	632,391	1,547	9,139,190
<b>80-84 years</b>	1,836	6,277,807	207	398,222	1,629	5,879,585
<b>85+ years</b>	2,083	6,721,134	236	351,802	1,847	6,369,333
<b>Total</b>	11,523	61,078,894	1,310	3,955,710	10,213	57,123,184

SOURCE: 2018 MCBS Internal Sample Control File

### 6.3 Primary and Secondary Sampling Units

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.

The MCBS selects its own PSUs and SSUs. Prior to Fall 2001, respondents for the MCBS were drawn from a sample of 107 PSUs that had been selected in 1991 from the 48 continental U.S. states, the District of Columbia, and Puerto Rico. A second-stage sample of 1,163 SSUs defined by ZIP Code was initially drawn within those PSUs. The second-stage sample was expanded each subsequent year to represent newly created ZIP Code areas, ultimately increasing to 1,523 SSUs in Fall 2000. For Fall 2001, the PSU sample was updated and reselected in a manner that maximized overlap with the original PSU sample. Within the new sample of 107 PSUs, 1,209 SSUs were initially selected in Fall 2001. With the addition of new ZIP Code clusters in subsequent years, the number of SSUs increased to 1,250 by Fall 2013.

Beginning in Fall 2014, census tracts or groups of tracts replaced ZIP Code areas as SSUs for the Incoming Panel selected each fall. A sample of 703 tract-based SSUs was selected within the existing 107 PSUs in 2014; the SSUs were sized to support beneficiary sampling for approximately 20 years. The new tract-based SSU design was chosen because census tracts are more stable and change less often than ZIP Code areas, resulting in less required maintenance. An additional benefit is that tract-based units are more easily merged to federal survey data such as those published by the Census (e.g., decennial census data and the American Community Survey (ACS)). While the MCBS PSUs and SSUs do not align directly with other surveys, they may overlap in some areas with PSUs and/or SSUs used for other federal health surveys. Respondents for the MCBS are sampled from the Medicare Administrative enrollment data. The MCBS sample is designed to yield about 14,500 completed cases annually in the MCBS Survey File and about 9,000 completed cases annually in the MCBS Cost Supplement File. Because of interest in their special health care needs, elderly beneficiaries (age 85 and over) and beneficiaries with disabilities (age 64 and under) are oversampled to permit more detailed analysis of these subpopulations.

Efforts were initiated in 2015 to more closely align our Hispanic oversample to the Hispanic population in the continental U.S., as most of our Hispanic beneficiaries were in Puerto Rico. These efforts began with the implementation of an additional Hispanic oversample within the continental U.S., with the intention of gradually reducing the selected sample in Puerto Rico and eventually ceasing to select sample from Puerto Rico altogether. However, the devastating hurricanes in Puerto Rico in 2017, which caused many beneficiaries to relocate, accelerated the removal of Puerto Rico from the sample frame. In 2017, the three PSUs in Puerto Rico were dropped from the original set of 107 PSUs, reducing the total MCBS PSUs to 104. Thus, the final 2018 Panel was selected from the remaining 104 MCBS PSUs, all of which are in the continental United States. At the second stage, the MCBS continued the use of the census tract-based secondary sampling units (SSUs) that were selected in 2014.<sup>14</sup> At the third stage, Medicare beneficiaries, the ultimate sampling units (USUs), were selected from within the selected tract-based SSUs.<sup>15</sup> In 2015, the sample design underwent significant modernization, including the aforementioned additional oversampling of Hispanic beneficiaries and the inclusion of current-year enrollees in the sampling frame; these changes are carried forward into 2018.

### *6.3.1 Eligibility: Medicare Population Covered by the 2018 LDS*

Beginning in 2015, beneficiaries who became eligible for Medicare Part A or B and enrolled anytime during the sampling year were eligible to be sampled as part of the annual panel. This is a substantial change in practice; prior to 2015, only beneficiaries enrolled in Medicare by January 1 of the sampling year were eligible to be sampled in an annual panel. More specifically, previously the MCBS would have waited until the 2016 Panel to select beneficiaries who became eligible and enrolled during 2015 (e.g., those 'new' to Medicare). Thus, to estimate 2015 events, cost, and utilization would require the 2015 Panel and all prior panels plus the new 2015 enrollees who were not sampled until the 2016 Panel. Beginning in 2015, these beneficiaries were selected as part of the 2015 Panel; thus, the 2015 Cost Supplement includes data using the 2015 Panel as well as the 2012, 2013, and 2014 previous panels, without the need to use data from the 2016 Panel. Likewise, the 2018 Cost Supplement includes data using the 2018 Panel as well as the 2015, 2016, and 2017 previous panels, without the need to use data from the 2019 Panel, allowing data to be released in a timelier manner. That is, data are released up to one year earlier with the Survey File LDS released 12-15 months after the end of data collection and the Cost Supplement LDS released 15-18 months after the end of data collection.

## **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.<sup>16</sup> 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 2015 through 2018 Panels.

<sup>14</sup> These SSUs were sized to last 20 years; therefore, there is no current plan to reselect the SSUs.

<sup>15</sup> In late 2000, the current set of PSUs was selected. In 2014, SSUs were reconstructed using Census tracts and a new sample was drawn. While the MCBS PSUs and SSUs do not align directly with other surveys, they may overlap in some areas with PSUs and/or SSUs used for other surveys.

<sup>16</sup> The MCBS 2018 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:** 2018 MCBS Sample Selection for the LDS Releases

Panel	# of Selected Beneficiaries	Previously Enrolled Beneficiaries	
		Still Alive as of January 1 of Panel Year	Newly Enrolled Beneficiaries Since Last Panel Selection
<b>2015</b>	8,621	Enrolled on or before 1/1/2014	Enrolled 1/2/2014 – 12/31/2015
<b>2016</b>	12,145	Enrolled before 1/1/2016	Enrolled 1/1/2016 – 12/31/2016
<b>2017</b>	11,623	Enrolled before 1/1/2017	Enrolled 1/1/2017 – 12/31/2017
<b>2018</b>	11,523	Enrolled before 1/1/2018	Enrolled 1/1/2018 – 12/31/2018

SOURCE: 2018 MCBS Internal Sample Control File

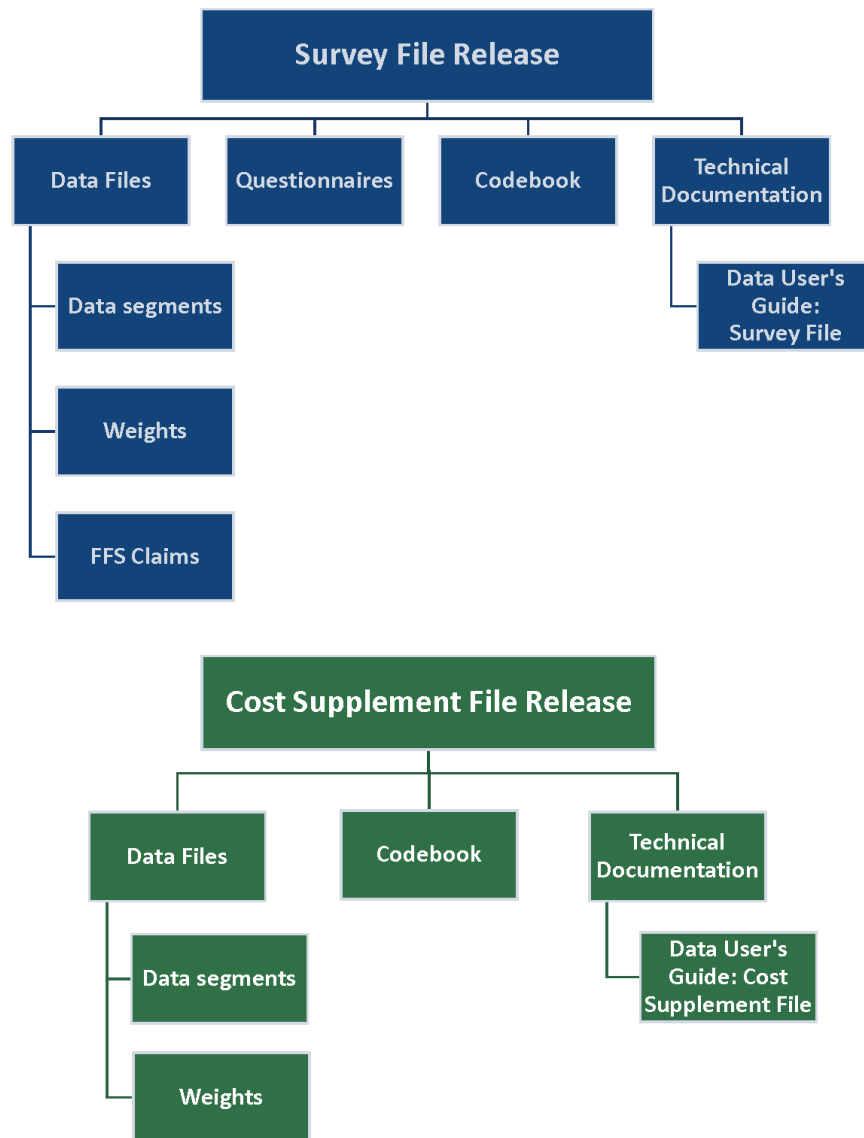
## 7. DATA PRODUCTS & DOCUMENTATION

### 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 one PUF (based on the Survey File data only) are planned. The LDS releases contain multiple files, called segments, which are easily linkable through a common beneficiary key ID. The Survey File LDS contains over 4,000 variables across 39 segments and the Cost Supplement LDS contains over 600 variables across 12 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.

**Exhibit 7.1:** 2018 Contents of Data Releases*7.1.1 2018 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. Exhibit 7.1.1 displays each segment included in the Survey File along with the abbreviation, description, and the equivalent historic segment from the 1991-2013 data release structure.

**Exhibit 7.1.1:** 2018 MCBS Survey File Segments and Contents

<b>Survey File Segment</b>	<b>Segment Abbrev</b>	<b>Description</b>	<b>Historic RIC Segment</b>	<b>Respondent Type (C, F, B)*</b>	<b>Data Source*</b>
<b>Access to Care</b>	ACCESSCR	Survey responses related to ability to obtain health care, delay of care related to costs, and reasons for not obtaining needed health care.	3	C	Community Questionnaire
<b>Access to Care Medical Appointment</b>	ACCSSMED	Survey responses related to medical and dental visit experiences and foregone medical, dental, vision, and hearing care. The ACCSSMED data that were collected in Winter 2019 are released with the 2018 Survey File. Special non-response adjustment weights are included with this file.	3	C	Community Questionnaire
<b>Administrative Utilization Summary</b>	ADMNUTLS	Summarized administrative information on Medicare, program expenditures, and utilization.	A	B	Administrative Records
<b>Assistance</b>	ASSIST	Identifies the person helping and type of assistance that the beneficiary may receive performing ADLs and IADLs (e.g., assistance with dressing, shopping, eating).	2H	C	Community Questionnaire
<b>Beneficiary Demographics</b>	DEMO	Demographic information collected in the survey and enhanced by Medicare Administrative data.	1, 9, A, K	B	Community Questionnaire, Facility Instrument, Administrative Records
<b>Chronic Conditions</b>	CHRNCOND	Survey responses related to chronic and other diagnosed medical conditions.	2, 2P	C	Community Questionnaire
<b>Chronic Conditions Flags</b>	CHRNCDL	FFS Chronic Condition Flag Records and FFS Chronic and other Disabling Flag records from administrative data sources.		B	Administrative Records
<b>Chronic Pain</b>	CHRNPAIN	Survey responses related to the beneficiary's experience with chronic pain and non-medication related chronic pain management techniques. The data collected in Summer 2019 are released with the 2018 Survey File. Special non-response adjustment weights are included with this file.	New	C	Community Questionnaire

<b>Survey File Segment</b>	<b>Segment Abbrev</b>	<b>Description</b>	<b>Historic RIC Segment</b>	<b>Respondent Type (C, F, B)*</b>	<b>Data Source+</b>
<b>Diabetes</b>	DIABETES	Survey responses related to diabetes management such as insulin usage.		C	Community Questionnaire
<b>Facility Assessments</b>	FACASMNT	Assessment information conducted while the beneficiary was living in a Medicare approved or non-Medicare approved facility.	2F	F	Facility Instrument
<b>Facility Characteristics</b>	FACCHAR	Primarily contains information from the Facility Questionnaire, while also incorporating Skilled Nursing Facility (SNF) stay information for beneficiaries living in the community and in facilities.	7, 7S	B	Facility Instrument
<b>Falls</b>	FALLS	Survey responses related to injuries and attitudes related to falls.	2, 2P	C	Community Questionnaire
<b>Food Insecurity</b>	FOODINS	Information regarding the beneficiary's access to sufficient food. The FOODINS data that were collected in Summer 2019 are released with the 2018 Survey File given that the reference period is 2018. Special non-response adjustment weights are included with this file.		C	Community Questionnaire
<b>General Health</b>	GENHLTH	Survey responses regarding a beneficiary's general health status and functioning such as height and weight.	2	C	Community Questionnaire
<b>Health Insurance Summary</b>	HISUMRY	Administrative information on the characteristics of insurance coverage.	4, A	B	Community Questionnaire, Administrative Records
<b>Health Insurance Timeline</b>	HITLINE	Types of insurance plans and the coverage eligibility timeline as well as information regarding premiums and covered services.	4, A	B	Community Questionnaire, Facility Instrument, Administrative Records
<b>Household (HH) Characteristics</b>	HHCHAR	Information about the beneficiary's household composition and home.	5	B	Community Questionnaire

Survey File Segment	Segment Abbrev	Description	Historic RIC Segment	Respondent Type (C, F, B)*	Data Source+
<b>Income and Assets</b>	INCASSET	Data on a beneficiary's income and assets. The INCASSET data were collected in Summer 2019 but released with the 2018 Survey File given that the reference period is 2018. Special non-response adjustment weights are included with this file.	1, Income Asset	B	Community Questionnaire
<b>Interview Characteristics</b>	INTERV	Summarizes the characteristics of the interview such as the type of interview conducted and whether a proxy was used.	4, 8, 9, K	B	Community Questionnaire, Facility Instrument
<b>Medicare Advantage (MA) Plan Questions</b>	MAPLANQX	Augments information from the Access to Care and Satisfaction with Care sections of the questionnaire for beneficiaries enrolled in Medicare Part C.	H	C	Community Questionnaire
<b>Medicare Plan Beneficiary Knowledge</b>	MCREPLNQ	Information about the beneficiary's experience with the Medicare open enrollment period and knowledge about Medicare covered expenses. The MCREPLNQ data were collected in Winter 2019 but released with the 2018 Survey File given that the reference period is 2018. Special non-response adjustment weights are included with this file.	KN	C	Community Questionnaire
<b>Minimum Data Set</b>	MDS3	Assessment information conducted while the beneficiary was living in an approved Medicare Facility.	MDS, 10	B	Administrative Records
<b>Mental Health</b>	MENTHLTH	Survey responses regarding the beneficiary's mental health such as feelings of anxiety or depression.		C	Community Questionnaire
<b>Mobility</b>	MOBILITY	Information on the beneficiary's use of available transportation options and whether the beneficiary's health affects their daily travel.		C	Community Questionnaire
<b>NAGI Disability</b>	NAGIDIS	Information on the beneficiary's difficulties with performance of activities of daily living.	2, 2H, 2P	C	Community Questionnaire
<b>Nicotine and Alcohol</b>	NICOALCO	Information on the prevalence and frequency of alcohol and nicotine use.	2, 2P	C	Community Questionnaire
<b>Outcome and Assessment Information</b>	OASIS	Assessment information conducted while the beneficiary was receiving home health services.	OAS, 10	B	Administrative Records

<b>Survey File Segment</b>	<b>Segment Abbrev</b>	<b>Description</b>	<b>Historic RIC Segment</b>	<b>Respondent Type (C, F, B)*</b>	<b>Data Source+</b>
<b>Patient Activation</b>	PNTACT	This questionnaire section is designed to assess the degree to which Medicare beneficiaries actively participate in their own health care and decisions concerning that care. Special non-response adjustment weights are included with this file.	PA	C	Community Questionnaire
<b>Preventive Care</b>	PREVCARE	Data on preventive services such as vaccinations and routine screening procedures.	2, 2P	C	Community Questionnaire
<b>RX Medications</b>	RXMED	Augments information from the Access to Care (ACQ) and Satisfaction with Care (SCQ) sections of the questionnaire regarding prescription medication access and satisfaction with and knowledge about Medicare Part D. The RXMED data were collected in Summer 2019, but released with the 2018 Survey File given that the reference period is 2018. Special non-response adjustment weights are included with this file.	RX	C	Community Questionnaire
<b>Residence Timeline</b>	RESTMLN	Information on where the beneficiary lived over the course of the year.	6, 9, A, K	B	Community Questionnaire, Facility Instrument
<b>Satisfaction with Care</b>	SATWCARE	Data on satisfaction with health care and reasons why beneficiaries do not seek medical care or prescription drugs.	3	C	Community Questionnaire
<b>Usual Source of Care</b>	USCARE	Data on where and how the beneficiary typically seeks medical care. The USCARE data were collected in Winter 2019, but released with the 2018 Survey File given that the reference period is 2018. Special non-response adjustment weights are included with this file.	2, 3	C	Community Questionnaire
<b>Vision and Hearing</b>	VISHEAR	Information on the beneficiary's eye health and hearing status.	2	C	Community Questionnaire

Survey File Segment	Segment Abbrev	Description	Historic RIC Segment	Respondent Type (C, F, B)*	Data Source†
<b>Weights</b>	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.	X, XE, X3, X4	B	Community Questionnaire, Facility Instrument
<b>Fee-for-Service Claims</b>	FFS	These files include 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).	Research Claims	B	Administrative Records

\* = Respondent type describes the expected setting where beneficiaries resided during the course of the calendar year (i.e., C = respondent only completed Community interviews, F = a Facility interview was conducted, or B = respondents completed at least one Community interview and for whom at least one Facility interview was conducted). 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).

† = Data source describes the source of the data on the segment. The three possible sources are the Community Questionnaire, Facility Instrument, and Administrative Records. Each LDS segment can have any combination of these sources. Data source is different from the Respondent type in that it reflects where the data came from, not where the respondent was living. For example, a beneficiary could have lived in both settings during the year so have a respondent type=B but the data for that beneficiary available on the ACCESSCR segment came from their Community interview only.

### 7.1.2 2018 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 healthcare 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.<sup>17</sup> 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.

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

<sup>17</sup> 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. In 2018, beneficiaries with MA or other private Medicare-approved health plans accounted for slightly over one in three Medicare beneficiaries (<https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Dashboard/Medicare-Enrollment/Enrollment%20Dashboard.html>).

healthcare 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.

**Exhibit 7.1.2:** 2018 MCBS Cost Supplement File Segments and Contents

<b>Cost Supplement Segment</b>	<b>Segment Abbrev</b>	<b>Description</b>	<b>Historic RIC Segment</b>	<b>Data Source<sup>+</sup></b>
<b>Dental Utilization Events</b>	DUE	This file contains individual dental events for the MCBS population.	DUE	Community Questionnaire, Administrative Records
<b>Facility Events</b>	FAE	This file includes individual facility events for the MCBS population. There is one record for each stay that occurred at least partly in the data year.	FAE	Facility Instrument, Administrative Records
<b>Inpatient Hospital Events</b>	IPE	This file contains individual inpatient hospital events for the MCBS population.	IPE	Community Questionnaire, Administrative Records
<b>Institutional Events</b>	IUE	This file contains individual short-term facility (usually SNF) stays for the MCBS population that were reported during a community interview or created from Medicare claims data.	IUE	Community Questionnaire, Facility Instrument, Administrative Records
<b>Medical Provider Events</b>	MPE	This file contains individual events for a variety of medical services, equipment, and supplies.	MPE	Community Questionnaire, Administrative Records
<b>Outpatient Hospital Events</b>	OPE	This file contains individual outpatient hospital events for the MCBS population.	OPE	Community Questionnaire, Administrative Records
<b>Prescribed Medicine Events</b>	PME	This file contains individual outpatient prescribed medicine events for the MCBS population.	PME	Community Questionnaire, Administrative Records
<b>Person Summary</b>	PS	Summarization of utilization and expenditures by type of service and summarization of expenditures by payer, yielding one record per person.	PS	Community Questionnaire, Facility Instrument, Administrative Records

Cost Supplement Segment	Segment Abbrev	Description	Historic RIC Segment	Data Source <sup>+</sup>
<b>Service Summary</b>	SS	Summarization of the seven individual event files along with home health and hospice utilization, yielding a total of nine summary records per person.	SS	Community Questionnaire, Facility Instrument, Administrative Records
<b>Cost Supplement Ever Enrolled Weights</b>	CSEVWGTS CSL2WGTS CSL3WGTS	The weights file provides: longitudinal weights for the ever enrolled population who had cost and utilization information, general-purpose cross-sectional weights, and a series of replicate weights.	X	N/A

<sup>+</sup> = Data source describes the source of the data on the segment. The three possible sources are the Community Questionnaire, Facility Instrument, and Administrative Records. Each LDS segment can have any combination of these sources.

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 analysts 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<sup>®</sup> formatted files and comma delimited files for use with Stata<sup>®</sup> and R<sup>®</sup>. Directions and sample SAS<sup>®</sup> code are given below to help users read the dataset into SAS<sup>®</sup> (see Appendix C). 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

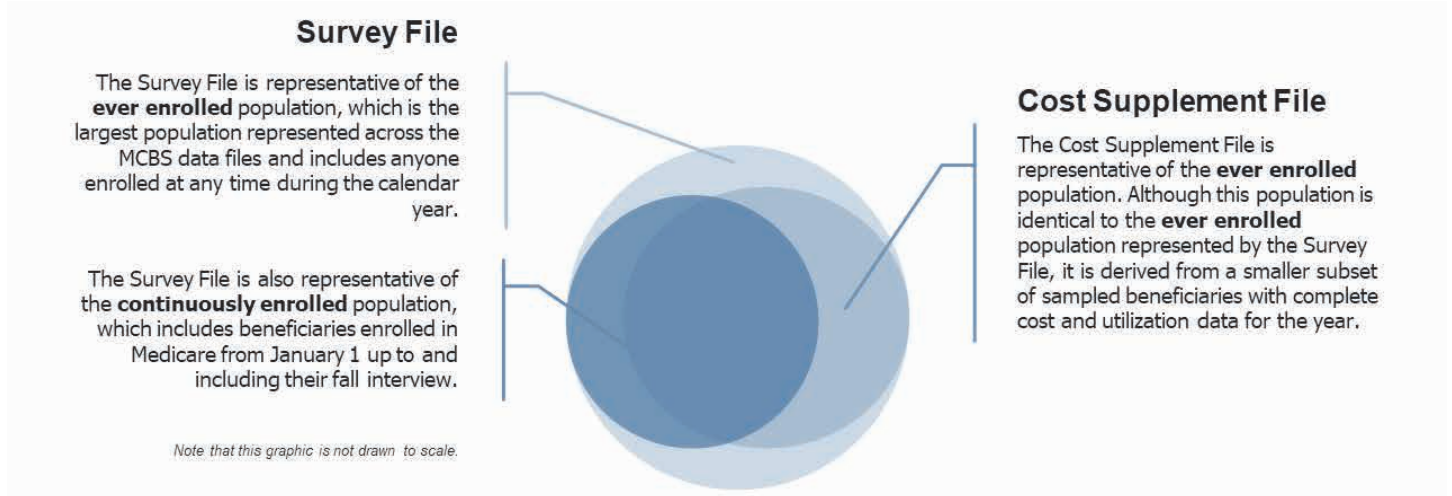
The fixed-length claims (also known as the research claims) are abbreviated versions of the full claim record layout. Each claim type has a subset of variables selected for their relevancy to data analysis of that service. Additionally, institutional claim types have a corresponding revenue center file 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.<sup>18</sup> The ever enrolled population from the Survey File is the largest, including anyone enrolled at any time during the calendar year. The continuously enrolled population is limited to those beneficiaries who were enrolled from January 1 of the survey year through the fall interview date. The Survey File includes weight segments that allow for subsetting the data by the ever enrolled and continuously enrolled populations. 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. Similarly, data users conducting year-to-year or longitudinal analyses with the 1991 through 2013 Access to Care files only need the Survey File to make comparisons of services.

### 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.

<sup>18</sup> Exhibit 7.2 is not drawn to scale, but provided as a visual reference for the relationship of populations between data files.

### 7.2.3 Using Both Community and Facility Data

To obtain an accurate representation of all Medicare beneficiaries, the MCBS sample includes all beneficiaries regardless of residence status. The MCBS follows beneficiaries into and out of long-term care facilities to maintain a comprehensive profile of their health care utilization and expenditures. About 5-8 percent of the sample are beneficiaries who live in a long-term care facility or who alternate between living in the community and living in a facility. 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.

Analytic decisions about whether to include all beneficiaries regardless of residence status, or those living only in the community or only in facilities, should be driven by both the research question and data limitations. However, since the underlying data collection approach differs based on the beneficiary's residence status, caution must 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 recommended steps should be followed:

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 10.9 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 limitations to the analysis.
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/Research-Statistics-Data-and-Systems/Research/MCBS/Data-Briefs>.

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

### 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 Fee-for-Service claims data. The claims records represent services provided during calendar year 2018 and processed by CMS. To facilitate analysis, the Administrative Utilization Summary files (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 "2018"), 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, an analyst 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.

### 8.3 Claims Files

The fixed-length claims (also known as the research claims or Fee-for-Service 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.

#### *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, 2018 and on or before December 31, 2018.
2. Skilled nursing facility claims were included if the admission or "from" date fell on or after January 1, 2018 and on or before December 31, 2018.
3. Home health agency and outpatient facility claims were included if the "through" date fell on or after January 1, 2018 and on or before December 31, 2018.
4. Hospice claims were included if the admission or "from" date fell on or after January 1, 2018 and on or before December 31, 2018.
5. Physician or supplier claims were included if the latest "service thru" date fell on or after January 1, 2018 and on or before December 31, 2018.
6. Durable medical equipment (DME) claims were included if the latest "service thru" date fell on or after January 1, 2018 and on or before December 31, 2018.

A total of 5,358 (about 35.2 percent) of the 2018 survey participants did not use Medicare reimbursed services in a FFS setting in 2018; 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.<sup>19</sup> For the other 9,879 individuals in the sample, the MCBS has captured claims meeting the date criteria, processed and made available by CMS through June 2019.<sup>20</sup> Medicare payment amounts have been reduced by the sequestration amount of two percent for all claims for service dates on or after April 1, 2013.

---

<sup>19</sup> The Health Insurance Timeline (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 Access to Care (ACCESSCR) and Access to Care, Medical Appointment (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.

<sup>20</sup> Note that claims "mature" through the midpoint of the following calendar year. That is, 2018 claims were pulled from CMS' administrative data after June 2019 to ensure that the 2018 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 those 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 Appointment 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 Questionnaire for each data year along with the questionnaire variable names and question text on the MCBS website at: <https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS/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](mailto: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 edits are conducted on the data files in order to check the data for accuracy, completeness, and reasonableness. Any structural issues are addressed during either data extraction or data cleaning.

Logic and reasonableness checks are also performed for each data file. Logic checks verify that the questionnaire worked as expected, particularly with respect to questionnaire routing. Errors identified during logic checking result in two categories of data edits: flagging values that were incorrectly skipped or setting incorrectly populated values to null to indicate a valid missing.

Global edits are applied to edit unreasonable or impossible extreme values to bind the data to reasonable responses and to check for values that are not explicitly disallowed by the questionnaire. For example, male respondents should not report female-only conditions, like cervical cancer. The MCBS also conduct consistency checks and edits. If a respondent reports becoming Medicaid eligible due to a certain condition, then they should have reported having that certain condition. Based on a thorough data review, these types of errors are corrected during data cleaning.

A flag variable is created for each edit to indicate whether the variable was edited for a particular observation for use in data processing quality control.

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, as 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**

In order 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.<sup>21</sup>

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 or not the corresponding value was imputed.

The 2018 Income and Assets imputation used IAQ data reported in 2019, as the 2019 IAQ asks about total income in the prior year (2018). The MCBS imputed different sets of variables for respondents to the 2019 IAQ and for the 2018 ever enrolled respondents who did not complete the 2019 IAQ. For the first group, the MCBS imputed a selection of variables from the 2019 IAQ. These included probe variables, which are indicators of whether the beneficiary and/or the spouse 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 had. For the second group, which includes beneficiaries living in a facility, only the amount of total income was imputed.

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.

<sup>21</sup> Kalton, Graham, and Daniel Kasprzyk. "The treatment of missing survey data." Survey methodology 12, no. 1 (1986): 1-16.

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.

For 2018, the Income and Assets imputation made greater use of survey reported income from earlier years, reducing the reliance on hot deck imputation based solely on demographic variables.

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 2018. For beneficiaries living in a facility, medical event data were provided only from Medicare claims data. The MCBS first imputed whether or not 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 or not the corresponding cost value was imputed.

## 9.4 Weighting

### *9.4.1 Preparing Statistics (Using the Full Sample Weights)*

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 2017 Survey File and Cost Supplement File. For analysis of Survey File data, there are two populations of inference that can be obtained through the use of 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, 2018, 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 2018.<sup>22</sup> 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 2018. To define the population, the MCBS creates a calendar history of a beneficiary's MCBS interviews. A number of eligibility checks are run against this calendar history to identify beneficiaries who met eligibility requirements for inclusion in the survey data for the calendar year, either because they were interviewed for a full year or interviewed until death or loss of Medicare entitlement. Beneficiaries who pass these eligibility checks become the population eligible for the Cost Supplement ever enrolled weight and the prescription medicine data files.

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.

<sup>22</sup> This is identical to the historical Access to Care (ATC) cross-sectional weight that was available in previous years, 1991-2013.

The weights mentioned above for the calendar year 2018 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 2018 Survey File and Cost Supplement LDS's.<sup>23</sup>

- **Survey File Two-Year Longitudinal Weights (LNG2WGTS):** Two-year longitudinal weights are populated only for members of the 2015, 2016, and 2017 panels who had 2017 and 2018 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 1/1/2017 and surviving and entitled as of completion of the Fall 2018 interview.
- **Survey File Three-Year Longitudinal Weights (LNG3WGTS):** Three-year longitudinal weights are populated only for members of the 2015 and 2016 panels who were continuously enrolled during all of the years 2016-2018 and had Survey File data in 2016 and 2018. The resulting weights represent the population of Medicare beneficiaries who enrolled on or before 1/1/2016 and were still alive and entitled as of completion of the Fall 2018 interview.
- **Survey File Four-Year Longitudinal Weights (LNG4WGTS):** Four-year longitudinal weights are populated only for members of the 2015 panel who were continuously enrolled during all of the years 2015-2018. The resulting weights represent the population of Medicare beneficiaries who enrolled on or before 1/1/2015 and were still alive and entitled as of completion of the Fall 2018 interview.
- **Cost Supplement Two-Year Longitudinal Weights (CSL2WGTS):** The two-year longitudinal weights are populated for members of the 2015, 2016, and 2017 panels who were ever enrolled in Medicare at any time during both 2017 and 2018 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 2015 and 2016 panels who were ever enrolled in Medicare at any time during 2016, 2017, and 2018, and provided utilization and cost data for all three years.

Topical Questionnaire sections related to the Survey File and Cost Supplement File are weighted separately as they are fielded in the winter and summer rounds following the data year, or are not administered to proxy respondents. There are three sets of full-sample and replicate weights for each module, one based on the 2018 Survey File ever enrolled population, one based on the 2018 Survey File continuously enrolled population, and one based on the 2018 Cost Supplement ever enrolled population. These weights may be used to conduct joint analyses of Topical data, Survey File data, and Cost Supplement 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. The Topical weights, segments, and weight names are listed in Exhibit 9.4.1. Please also see section 10.37 for information on using the Topical weights and the forthcoming *2018 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.

<sup>23</sup> 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.

**Exhibit 9.4.1:** 2018 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 1/1/2018 through the fall of 2018
Survey File	Ever Enrolled Cross-Sectional Weights	EVRWGTS	EEYRSWGT	EEYRS001-EEYRS100	Ever enrolled for at least one day at any time during 2018
Survey File	Continuously Enrolled Two-Year Longitudinal Weights	LNG2WGTS	L2YRSWGT	L2YRS001-L2YRS100	Continuously enrolled from 1/1/2017 through the fall of 2018
Survey File	Continuously Enrolled Three-Year Longitudinal Weights	LNG3WGTS	L3YRSWGT	L3YRS001-L3YRS100	Continuously enrolled from 1/1/2016 through the fall of 2018
Survey File	Continuously Enrolled Four-Year Longitudinal Weights	LNG4WGTS	L4YRSWGT	L4YRS001-L4YRS100	Continuously enrolled from 1/1/2015 through the fall of 2018
Cost Supplement File	Ever Enrolled Cross-Sectional Weights	CSEVRWGT	CSEVRWGT	CSEVR001-CSEVR100	Ever enrolled for at least one day at any time during 2018
Cost Supplement File	Two-Year Longitudinal Weights	CSL2WGTS	CSL2YWGT	CSL2Y001-CSL2Y100	Enrolled at any time during both 2017 and 2018
Cost Supplement File	Three-Year Longitudinal Weights	CSL3WGTS	CSL3YWGT	CSL3Y001-CSL3Y100	Enrolled at any time during each of 2016, 2017, and 2018
Survey File Topical Section	KNQ Survey File Ever Enrolled	MCREPLNQ	KNSEWT	KNSE1-KNSE100	Ever enrolled in 2018 and still alive, entitled, and not residing in a facility in Winter 2019

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 2018 and still alive, entitled, and not residing in a facility in Winter 2019
Survey File Topical Section	KNQ Cost Supplement Ever Enrolled	MCREPLNQ	KNCEWT	KNCE1-KNCE100	Ever enrolled in 2018 and still alive, entitled, and not residing in a facility in Winter 2019
Survey File Topical Section	ACQ Survey File Ever Enrolled	ACCSSMED	ACSEWT	ACSE1-ACSE100	Ever enrolled in 2018 and still alive, entitled, and not residing in a facility in Winter 2019
Survey File Topical Section	ACQ Survey File Continuously Enrolled	ACCSSMED	ACSCWT	ACSC1-ACSC100	Continuously enrolled in 2018 and still alive, entitled, and not residing in a facility in Winter 2019
Survey File Topical Section	ACQ Cost Supplement Ever Enrolled	ACCSSMED	ACCEWT	ACCE1-ACCE100	Ever enrolled in 2018 and still alive, entitled, and not residing in a facility in Winter 2019
Survey File Topical Section	USQ Survey File Ever Enrolled	USCARE	USSEWT	USSE1-USSE100	Ever enrolled in 2018 and still alive, entitled, and not residing in a facility in Winter 2019
Survey File Topical Section	USQ Survey File Continuously Enrolled	USCARE	USSCWT	USSC1-USSC100	Continuously enrolled in 2018 and still alive, entitled, and not residing in a facility in Winter 2019

Limited Data Set	Description	Segment	Full-Sample Weight	Replicate Weights	Population
Survey File Topical Section	USQ Cost Supplement Ever Enrolled	USCARE	USCEWT	USCE1-USCE100	Ever enrolled in 2018 and still alive, entitled, and not residing in a facility in Winter 2019
Survey File Topical Section	IAQ Survey File Ever Enrolled	INCASSET	INSEWT	INSE1-INSE100	Ever enrolled in 2018 and still alive, entitled, and not residing in a facility in Summer 2019
Survey File Topical Section	IAQ Survey File Continuously Enrolled	INCASSET	INSCWT	INSC1-INSC100	Continuously enrolled in 2018 and still alive, entitled, and not residing in a facility in Summer 2019
Survey File Topical Section	IAQ Cost Supplement Ever Enrolled	INCASSET	INCEWT	INCE1-INCE100	Ever enrolled in 2018 and still alive, entitled, and not residing in a facility in Summer 2019
Survey File Topical Section	IAQ Survey File Ever Enrolled	FOODINS	FDSEWT	FDSE1-FDSE100	Ever enrolled in 2018 and still alive, entitled, and not residing in a facility in Summer 2019
Survey File Topical Section	IAQ Survey File Continuously Enrolled	FOODINS	FDSCWT	FDSC1-FDSC100	Continuously enrolled in 2018 and still alive, entitled, and not residing in a facility in Summer 2019
Survey File Topical Section	IAQ Cost Supplement Ever Enrolled	FOODINS	FDCEWT	FDCE1-FDCE100	Ever enrolled in 2018 and still alive, entitled, and not residing in a facility in Summer 2019

Limited Data Set	Description	Segment	Full-Sample Weight	Replicate Weights	Population
Survey File Topical Section	PAQ Survey File Enrolled	PNTACT	PASEWT	PASE1-PASE100	Ever enrolled for at least one day at any time during 2018
Survey File Topical Section	PAQ Survey File Continuously Enrolled	PNTACT	PASCWT	PASC1-PASC100	Continuously enrolled from 1/1/2018 through the fall of 2019
Survey File Topical Section	PAQ Cost Supplement Ever Enrolled	PNTACT	PACEWT	PACE1-PACE100	Ever enrolled for at least one day at any time during 2018
Survey File Topical Section	RXQ Survey File Ever Enrolled	RXMED	RXSEWT	RXSE1-RXSE100	Ever enrolled in 2018 and still alive, entitled, and not residing in a facility in Summer 2019
Survey File Topical Section	RXQ Survey File Continuously Enrolled	RXMED	RXSCWT	RXSC1-RXSC100	Continuously enrolled in 2018 and still alive, entitled, and not residing in a facility in Summer 2019
Survey File Topical Section	RXQ Cost Supplement Ever Enrolled	RXMED	RXCEWT	RXCE1-RXCE100	Ever enrolled in 2018 and still alive, entitled, and not residing in a facility in Summer 2019
Survey File Topical Section	CPQ Survey File Ever Enrolled	CHRNPAIN	CPSEWT	CPSE1-CPSE100	Ever enrolled in 2018 and still alive, entitled, and not residing in a facility in Summer 2019
Survey File Topical Section	CPQ Survey File Continuously Enrolled	CHRNPAIN	CPSCWT	CPSC1-CPSC100	Continuously enrolled in 2018 and still alive, entitled, and not residing in a facility in Summer 2019

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 2018 and still alive, entitled, and not residing in a facility in Summer 2019

## 9.5 Using the Data

### 9.5.1 Merging Segments within 2018

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<sup>®</sup> code for the construction of an analytic file using the 2018 Survey File LDS. For an example of how to merge data across the Survey File and Cost Supplement File LDS's, please see Appendix C.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) 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. Analysts 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.<sup>24</sup>

To estimate variance using the balanced repeated replication method, a series of replicate weights are included in the 2018 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

<sup>24</sup> Wolter, Kirk. Introduction to variance estimation. Springer Science & Business Media, 2007.

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. For examples and guidance on using the Taylor Series linearization method of variance estimation or the BRR method, please see Appendix C.

### *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<sup>®</sup>, STATA<sup>®</sup>, SUDAAN<sup>®</sup>, and the complex survey procedures in SAS<sup>®</sup>.

### *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 age 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 analyst 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 C 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*

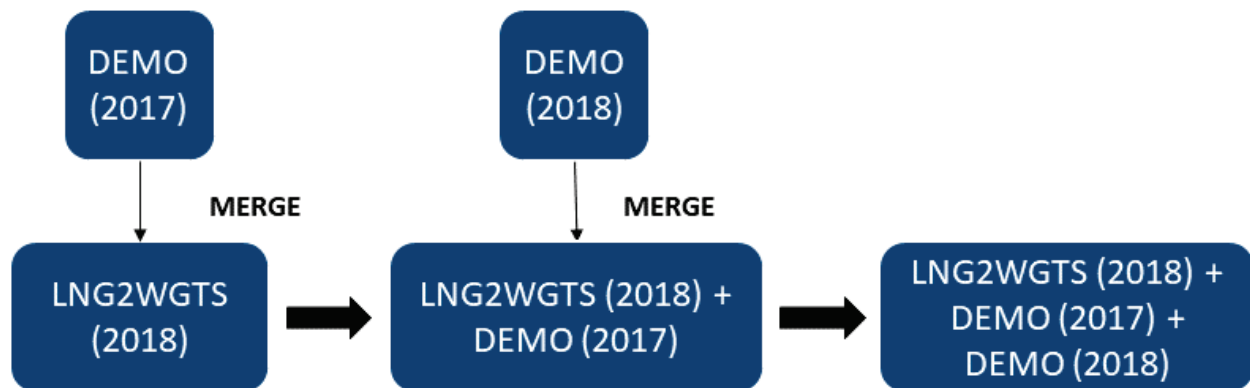
The study objective in longitudinal analysis is to assess changes over time for each sample person. The Survey File cross-sectional and longitudinal population definitions are consistent from year to year, so the data are

comparable between years. The Cost Supplement cross-sectional population definition is also consistent and comparable from year to year.<sup>25</sup>

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 CY2017 to CY2018 – 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 “backward 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 “backward 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 “backward longitudinal weights,” there are corresponding sets of replicate weights. The replicate weights included in the “backward 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 “backward longitudinal weights” file (LNG2WGTS). The second set of replicate longitudinal weights is labeled L3YRS001 through L3YRS100 and may be found on the three-year “backward 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 “backward longitudinal weights” file (LNG4WGTS).

<sup>25</sup> The Cost Supplement two-year longitudinal population changed slightly in 2016 from what was defined the last time the two-year longitudinal weights were supplied (i.e., in 2013). In 2013, the two-year longitudinal (i.e., one-year backward longitudinal weight) Cost Supplement weights represented the population that enrolled on or before 1/1/2011 and were still enrolled in 2013 (i.e., enrollees after 1/1/2011 were not included). Beginning in 2016, the two-year longitudinal weights represent a true two-year ever enrolled population (i.e., the population of beneficiaries that were ever enrolled in both 2015 and 2016).

### 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. 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 CY2017 and CY2018 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. 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).

## 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 participants) or to globally edit some variables. Variables preceded by the character "H\_", such as H\_DOB, come from CMS administrative source files.

#### *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 D.

#### *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 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 the fall round in certain scenarios. 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 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.

For more information about "ever" variables pertaining to chronic conditions, see the data notes for the Chronic Conditions segment (CHRNCOND) in Section 10.7.

### *10.1.6 Data Editing*

Data are edited for consistency and to provide users with files that are easily used for analysis.

### *10.1.7 Other Specify Questions*

Respondents are asked a number of questions with "other specify" response options. For example, respondents are asked about types of problems they may have experienced in getting health care. If respondents provide answers that are not on the list of possible choices, interviewers select the option for "other specify," in addition to recording their actual responses verbatim.

Beginning with the 2016 data, a new programmatic backcoding approach was developed for some "other specify" variables to streamline the review and categorization of verbatim responses. The 2016 backcoding implementation involved programmatically fixing misspellings of commonly used terms and then categorizing responses into existing response option categories as appropriate using keyword searches for terms matching those categories. Codes are then assigned to similar responses to facilitate analysis; there are no verbatim responses provided on the released segments. 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 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.

### *10.1.9 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 tell you if the information is survey-reported only, administrative data only, or both. Data that are only administrative are indicated as such in the data documentation and codebook.

## **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 Access to Care (ACCESSCR)**

### *10.3.1 Core Content*

The Access to Care segment contains information from the Health Status and Functioning (HFQ) section of the Fall 2018 questionnaire. General questions were asked about the beneficiary's ability to access medical services in 2018. This segment also contains information on medical debt and the reasons beneficiaries could not access the care they needed.

### *10.3.2 Variable Definitions*

CGET variables (e.g., CGETAPPT, CGETAPRV): The verbatim question OFFEXVB1 is back coded as necessary into variables that explain why a doctor's office indicated that it is difficult for Medicare patients to get an appointment/Medicare is not accepted, but the verbatim text is not released.

### *10.3.3 Special Notes*

Respondents were asked several questions 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 answered these questions in their own words, and interviewers selected 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.

## 10.4 Access to Care, Medical Appointment (ACCSSMED)

### 10.4.1 Core Content

The Access to Care segment contains information from the Access to Care (ACQ) section of the questionnaire from Winter 2019 and from the emergency room, outpatient, medical provider, and dental utilization sections asked in Winter 2019. General questions were asked about the beneficiary's access to all types of medical services in 2018 and the reasons for their visits or why they did not have visits for particular types of medical care. Questions about specific types of medical visit experiences were asked in context of the utilization section of the questionnaire in Winter 2019 so all respondents would be asked the questions in the same manner.

### 10.4.2 Variable Definitions

**D\_ERTIME:** The length of time the beneficiary spent waiting in the hospital emergency room before they saw a doctor or some other medical person is collected in either hours, minutes, or some combination of hours and minutes. The length of time is stored in a continuous variable while the corresponding unit is stored in a categorical variable (hours, minutes, or hours and minutes). These variables are used to derive D\_ERTIME, which is the length of time spent waiting in the hospital emergency room in minutes.

**D\_MDAPPT:** The length of time the beneficiary spent waiting for an appointment with a medical doctor is collected in days, weeks, or months. The length of time is stored in a continuous variable while the corresponding unit is stored in a categorical variable (days, weeks, or months). These variables are used to derive D\_MDAPPT, which is the length of time spent waiting for an appointment with a medical doctor in days.

**D\_MDTIME:** The length of time the beneficiary spent waiting during their most recent visit to see a doctor or a medical person (including time spent in the waiting room or exam room) is collected in either hours, minutes, or some combination of hours and minutes. The length of time is stored in a continuous variable while the corresponding unit is stored in a categorical variable (hours, minutes, or hours and minutes). These variables are used to derive D\_MDTIME, which is the length of time spent waiting to see a doctor or a medical person in minutes.

**D\_OPAPPT:** The length of time the beneficiary spent waiting for an appointment with a hospital clinic or outpatient department collected in days, weeks, or months. The length of time is stored in a continuous variable while the corresponding unit is stored in a categorical variable (days, weeks, or months). These variables are used to derive D\_OPAPPT, which is the length of time spent waiting for an appointment with a hospital clinic or outpatient department in days.

**D\_OPTIME:** The length of time the beneficiary spent waiting during their most recent visit to a hospital clinic or outpatient department to see a doctor or a medical person (including time spent in the waiting room or exam room) is collected in either hours, minutes, or some combination of hours and minutes. The length of time is stored in a continuous variable while the corresponding unit is stored in a categorical variable (hours, minutes, or hours and minutes). These variables are used to derive D\_OPTIME, which is the length of time spent waiting to see a doctor or a medical person at a hospital clinic or outpatient department in minutes.

### 10.4.3 Special Notes

In 2017, the DENTAL segment was eliminated and all items moved to the ACCSSMED segment with other questions regarding access to health care. Analyses of these items should use the Topical weights included with the segment. In 2018, several items from the Health Status and Functioning (HFQ) on delayed care for hearing and vision services were added to the segment. Please see exhibit 3.5.1 for variable names and descriptions.

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 the Weights section 10.36.3 for information on using weights with data from Topical questionnaire sections.

## 10.5 Administrative Utilization Summary (ADMNUTLS)

### 10.5.1 Core Content

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

### 10.5.2 Variable Definitions

Except as noted otherwise, the variables in this segment were derived from summarizing data from CMS's Medicare Administrative enrollment data and the Medicare Administrative utilization and payment records. Administrative data available as of December 31, 2018 were summarized to create these data items.

**H\_HHASW:** One or more home health agency 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\_DMEEVT, H\_DMEPMT, H\_TST EVT, H\_TSTPMT, H\_ANEVT, H\_ANEPMT, H\_ASCEVT, H\_ASCPMT, H\_DIAEVT, H\_DIAPMT, H\_EMEVT, H\_EMPMT, H\_IMG EVT, H\_IMG PMT, 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\_SNFSY, H\_SNF DAY. 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 was selected for each claim type in preparing this calculation. The CLM\_VAL\_CD = '64' was 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 was selected for each claim type in preparing

this calculation. The CLM\_VAL\_CD = '65' was 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 for a given 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 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\_ACTSTY:** Acute Inpatient Stays is the count of acute inpatient hospital stays (unique admissions, which may span more than one facility) for the 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 year.

**H\_ACTBPT:** Acute Inpatient Beneficiary Payments is the sum of Medicare coinsurance and deductible payments in the acute inpatient hospital setting for the 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  $\geq 0$ .

**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 a given 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 claims in 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 other inpatient (OIP) setting for the 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\_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\_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 a given 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 year. This variable equals the sum of the CLM\_UTLZTN\_DAY\_CNT variables on the source claims. 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\_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 "other" inpatient (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  $\geq 0$ . These OIP claims are a subset of the claims in the IP data file 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\_SNFPMT:** SNF Medicare Payments is the total Medicare payments in the SNF setting for the year.

**H\_SNFBPT:** SNF Medicare Payments is the total Medicare payments in the SNF setting for the year.

**H\_SNFSTY:** SNF Stays is the count of SNF stays (unique admissions, which may span more than one facility) for a given 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\_SNFDAY:** SNF Medicare Covered Days is the count of Medicare covered days in the SNF setting for the year. This variable equals the sum of the CLM\_UTLZTN\_DAY\_CNT variables on the source claims.

**H\_SNFBPT:** Skilled Nursing Facility Beneficiary Payments is the sum of Medicare coinsurance and deductible payments in the skilled nursing facility (SNF) setting for the year. The total beneficiary payment is calculated as the sum of the beneficiary deductible amount and coinsurance amount (variables called NCH\_BENE\_IP\_DDCTBL\_AMT and NCH\_BENE\_PTA\_COINSRNC\_LBLTY\_AM) for all SNF claims where the CLM\_PMT\_AMT  $\geq 0$ .

**H\_HOSPSTY:** Hospice Stays is the count of stays (unique admissions, which may span more than one facility) in the hospice setting for a given year. A hospice 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 hospice setting for a given year. This variable equals the sum of the CLM\_UTLZTN\_DAY\_CNT variables on the source claims.

**H\_HOSDAY:** Hospice Medicare Covered Days is the count of Medicare covered days in the hospice setting for a given 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 a given year.

**H\_HHVIS:** Home Health Visits is the count of home health (HH) visits for the year.

**H\_HOPPMT:** Hospital Outpatient Medicare Payments is the total Medicare payments in the hospital outpatient (HOP) setting for a given 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 year.

**H\_HOP\_ER:** Hospital Outpatient Emergency Rm Visits is the count of unique emergency department revenue center dates (as a proxy for an ED visit) in the hospital outpatient claims for the year. Revenue center codes indicating Emergency Room use were (0450, 0451, 0452, 0456, or 0459).

**H\_HOBPT:** Hospital Outpatient Beneficiary Payments is the sum of Medicare coinsurance and deductible payments in the hospital outpatient (HOP) setting for a given year. The total beneficiary payment is calculated as the sum of the beneficiary deductible amount and coinsurance amount (variables called `REV_CNTR_CASH_DDCTBLE_AMT` and `REV_CNTR_COINSRNC_WGE_ADJSTD_C`) for all HOP claims where the `CLM_PMT_AMT`  $\geq 0$ .

**H\_PB\_DRG:** Part B Drug Medicare Payments is the total Medicare payments for Part B drugs for a given 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 a given year. An event is defined as each line item that contains the relevant service. Part B drug claims are a subset of the claims in the Part B Carrier and DME claims.

**H\_BPTDRG:** Part B Drug Beneficiary Payments is the sum of coinsurance and deductible payments for part B drugs for a given year. The total Beneficiary payments are calculated as the sum of `LINE_COINSRNC_AMT` + `LINE_BENE_PTB_DDCTBL_AMT` for the relevant lines. Part B drug claims are a subset of the claims in the Part B Carrier and DME data files. 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 evaluation and management services for a given 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 evaluation and management (E&M) services for a given year. The total Beneficiary payments are calculated as the sum of `LINE_COINSRNC_AMT` + `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 a given year. Physician office 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 in the part B physician office services (PHYS) for a given year. An event is defined as each line item that contains the relevant service. Physician office 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\_PHYBPT:** Part B Physician Beneficiary Payments is the sum of coinsurance and deductible payments for the part B physician office services (PHYS) for a given year. The total Beneficiary payments are calculated as the sum of `LINE_COINSRNC_AMT` + `LINE_BENE_PTB_DDCTBL_AMT` for the relevant lines. Physician office claims

are a subset of the claims in the Part B Carrier and DME data files, and a subset of physician evaluation and management claims (note that E&M are tabulated separately in this data file). 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 a given 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 a given 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 (i.e., not anesthesia or dialysis) for a given year. The total Beneficiary payments are calculated as the sum of LINE\_COINSRNC\_AMT + 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 a given 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 durable medical equipment (DME) for a given year. An event is defined as each line item that contains the relevant service. Claims for DME are a subset of the claims in the Part B Carrier and DME claims.

H\_DMEBPT: Durable Medical Equipment Beneficiary Payments is the total Medicare payments for part B durable medical equipment (DME) for a given 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 a given 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 (OTH) 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 a given year, which includes Part B Carrier and DME claims which appear in specific settings for a given 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 (OTH) include ambulance, chiropractor, chemotherapy, vision, hearing and speech services, etc. 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 which appear in settings other than the 10 specific categories which are part of this file for a given year. The total Beneficiary payments are calculated as the sum of LINE\_COINSRNC\_AMT + LINE\_BENE\_PTB\_DDCTBL\_AMT for the relevant lines. Claims for other carrier/DME claims are a subset of the claims in the Part B Carrier and DME data files. Types of services which may have been summarized in this other carrier category (OTH) include ambulance, chiropractor, chemotherapy, vision, hearing and speech services, etc.

**H\_PTDPMT:** Part D Medicare Payments is the dollar amount that the Part D plan covered for all covered drugs for a given 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 a given year (i.e., a unique count of the PDE\_IDs). An event is a dispensed (filled) drug prescription that appears on the source Prescription Drug Event (PDE) claims.

**H\_PTDBPT:** Part D Beneficiary Payments is the dollar amount that the beneficiary paid for all PDEs for a given 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 Prescription Drug Events (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 a given year. This value includes the ingredient cost, dispensing fee, sales tax (if applicable), and vaccine administration fee.

**H\_ASCEVT:** This is the count of events in the part B ambulatory surgery center setting for a given 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 ambulatory surgery center (ASC) setting for a given year. The total beneficiary payment is calculated as the sum of the LINE\_COINSRNC\_AMT + 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:** This is the total Medicare payments for part B anesthesia services for a given year. Anesthesia claims are a subset of the claims, and a subset of procedures in the Part B Carrier claims.

**H\_ANEVT:** This is the count of events for part B anesthesia services for a given 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 anesthesia services (ANES) for a given year. The total Beneficiary payments are calculated as the sum of LINE\_COINSRNC\_AMT + 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:** This is the total Medicare payments for Part B dialysis services (primarily the professional component since treatments are covered in hospital outpatient) for a given year. Dialysis claims are a subset of the claims, and a subset of procedures in the Part B Carrier claims.

**H\_DIAEVT:** This is the total Medicare payments for Part B dialysis services (primarily the professional component since treatments are covered in hospital outpatient) for a given year. An event is defined as each line item that contains the relevant service. Dialysis claims are a subset of the claims, and a subset of procedures in the Part B Carrier claims.

**H\_DIABPT:** Dialysis Beneficiary Payments is the total Medicare payments for Part B dialysis services (primarily the professional component since treatments are covered in hospital outpatient) for a given year. The total Beneficiary payments are calculated as the sum of LINE\_COINSRNC\_AMT + LINE\_BENE\_PTB\_DDCTBL\_AMT for the relevant lines.

**H\_IMGPMT:** This is the total Medicare payments for imaging services for a given 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\_IMGEVT:** This is the count of events for imaging services for a given year. An event is defined as each line item that contains the relevant service. Claims for imaging procedures are a subset of the claims, and a subset of procedures in the Part B Carrier and DME claims.

**H\_IMGBPT:** Imaging Beneficiary Payments is the sum of coinsurance and deductible payments for imaging services (IMG) for a given year. The total beneficiary payments are calculated as the sum of `LINE_COINSRNC_AMT + LINE_BENE_PTB_DDCTBL_AMT` for the relevant lines. Claims for imaging procedures are a subset of the claims, and a subset of procedures in the Part B Carrier and DME data files. These imaging claims are defined as those with a line BETOS code (BETOS\_CD) where the first digit = 1 (except for '11E', or '11F' – which are considered Part B drugs).

**H\_TSTPMT:** This is the total Medicare payments for part B tests for a given year. Claims for tests are a subset of the claims in the Part B Carrier claims.

**H\_TSTEVT:** This is the count of events for part B tests for a given 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 a given 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).

### 10.5.3 Special Notes

For easier comparison of groups of people 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 2018, as received and processed by CMS through December 2019 for the 2018 benefit year. The administrative data source for this information changed in 2016. There are different breakouts and summary items than on previous versions of MCBS data.

The utilization summary represents services rendered and reimbursed under Medicare FFS in the calendar year 2018. 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 empty/missing. If the beneficiary used no services of a particular type (e.g., inpatient hospitalization), the variables relating to those benefits will be empty/missing.

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

## 10.6 Assistance (ASSIST)

### 10.6.1 Core Content

This segment contains information on each person identified as helping the beneficiary to perform 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 ASSIST segment reflects the number of persons identified as having assisted the beneficiary in performing one or more ADL or IADLs. Therefore, it is possible to have one, several, or no records per beneficiary.

### *10.6.2 Variable Definitions*

**HLPRNUM:** This variable is the Helper Identification Number and is derived from the survey's administrative files. The survey develops a person roster containing information about each person living with, treating, or helping the beneficiary.

**HLPRMOST:** When a beneficiary has more than one helper, this variable identifies which helper provides the beneficiary with the most help with daily activities. This variable is coded as a 1 for the helper who gave the most help, and is missing for all other helpers for that beneficiary. This variable also 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 all of that beneficiary's helpers.

### *10.6.3 Special Notes*

During final 2017 data processing, it was discovered that rostering information for newly rostered helpers was incorrectly excluded. This resulted in missing records in the ASSIST segment and missing data for the variable HLPREL in some records in the ASSIST segment in data years 2016, 2017, and 2018. This issue was resolved in the questionnaire in fall 2018.

## **10.7 Chronic Conditions (CHRNCOND)**

### *10.7.1 Core Content*

The Chronic Conditions segment provides data on whether the beneficiary had a series of chronic and other diagnosed medical conditions such as cancer, high blood pressure, and depression. If the beneficiary responds that they have the condition, a series of follow-up questions is asked.

### *10.7.2 Variable Definitions*

**Note:** The answers in the HFQ section of the questionnaire reflect the respondent's opinion, not a professional medical opinion.

**BLOOD PRESSURE:** A number of variables asking about blood pressure appear in this segment (e.g., the beneficiary's age when they first learned of their high blood pressure, whether the beneficiary measures their blood pressure at home, whether the beneficiary takes medication for high blood pressure, etc.).

**D\_OCDTYP:** This variable, indicating type of diabetes, is derived from HFQ items OCDTYPE and DIAPRGNT. The OCDTYPE categories for "Pre-diabetes" and "Borderline" diabetes are combined into one category for D\_OCDTYP. Female beneficiaries who answered "Yes" for DIAPRGNT, which is not released, are coded as "Gestational diabetes" for D\_OCDTYP, unless they indicated for OCDTYPE that they had Type 1 diabetes.

**HYSTEREC:** Female respondents were asked if they have ever had a hysterectomy in the last year. "Hysterectomy" includes partial hysterectomies. This variable is not asked of:

- male beneficiaries
- female beneficiaries in the Incoming Panel sample other than those who reported that they have never had a hysterectomy

- female beneficiaries in the Continuing sample who previously reported having had a hysterectomy in an earlier round

**ILLNESS/CONDITION VARIABLES:** The MCBS asks respondents whether they have ever had any of a series of illnesses or conditions. Their responses are coded affirmatively if the respondent had at some time been diagnosed with the conditions, even if the condition had been corrected by time or treatment. The condition must have been reported by the respondent as diagnosed by a physician, and not by the respondent. If the respondent was not sure about the definition of a condition, the interviewer offered no advice or information, but recorded the respondent's answer verbatim. The MCBS asks about: heart disease and high blood pressure; disorders or diseases of the brain; psychiatric disorders; intellectual disability; skin cancer; cancer, other than skin cancer; diabetes; arthritis; osteoporosis; a broken hip; emphysema, asthma, or chronic obstructive pulmonary disease (COPD); complete or partial paralysis; an amputation; enlarged prostate or benign prostatic hypertrophy.

If the respondent confirms having had cancer, other than skin cancer, a series of follow-up questions is asked to identify the kind of cancer.

In the fall round, all respondents are asked about various illnesses or conditions, such as hypertension. There are different versions of each 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 (hypertension, for example). 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 variables 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 were diagnosed with that condition in the past year irrespective of prior responses. All data from a beneficiary from the current survey year and all previous years are used to determine whether the beneficiary had 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.

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

### *10.7.3 Special Notes*

Two new variables were added to CHRNCOND for the 2018 data year: BLOSWGHT ("To lower risk of certain disease, has a doctor or health professional ever told you to lose weight?") and BASKDPRS ("Has a doctor or health professional ever asked you if there was ever a period of time when you felt sad, empty, or depressed?").

## **10.8 Chronic Condition Flags (CHRNCDL)**

### *10.8.1 Core Content*

Beginning in 2017, Chronic Conditions Flags and Chronic and Other Disabling Conditions flags from administrative Fee-for-Service (FFS) records are included. These flags are taken from administrative data from the Chronic Conditions Warehouse (CCW) which summarizes the beneficiaries FFS claims for the calendar year, and provides whether or not 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 particular 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.

### *10.8.2 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 also is 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 will be missing for beneficiaries that have never had the condition. The earliest possible date for anyone is January 1, 1999. If the beneficiary became eligible for Medicare after that, the earliest possible date will be some time after his/her coverage start date.

### *10.8.3 Special Notes*

These data are pulled from the CCW data sources. The criteria were 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>.

## **10.9 Chronic Pain (CHRNPAIN)**

### *10.9.1 Core Content*

Beginning in 2018, data on beneficiaries' experiences with chronic pain and chronic pain management techniques were collected in the summer round of 2019 in the Topical Chronic Pain Questionnaire (CPQ). 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). CPQ data are released in the CHRNPAIN segment.

### *10.9.2 Variable Definitions*

**PAINOFTN:** Indicates beneficiary's frequency of pain (never, some days, most days, or every day).

**PAINAMNT:** For beneficiaries who had any pain, this variable indicates the amount of pain (a little, a lot, or somewhere between a little and a lot).

**PAINLIMT and PAINFAM:** These two variables indicate how frequently chronic pain limits what the beneficiaries are able to do (never, some days, most days, or every day). **PAINLIMT** reflects the extent to which pain limits everyday life and work activities, and **PAINFAM** indicates the extent to which pain affects family and significant others.

**PAINMANG:** Reflects whether beneficiaries are able to manage pain to do things that they enjoy (not at all, a little, a lot, or somewhere between a little and a lot).

**MANGPHYS, MANGSPNE, MANGCBT, MANGPROG, MANGGROP, MANGYOGA, MANGMASG, and MANGMEDT:** These variables reflect whether beneficiaries are engaged in certain non-medication pain management techniques (yes or no).

**PAINBACK, PAINARMS, PAINLEGS, PAINHEAD, PAINABDM, and PAINOTOTH:** These variables indicate how much pain beneficiaries experienced in specific locations (not at all, a little, a lot, or somewhere between a little and a lot).

### *10.9.3 Special Notes*

The CPQ is administered the summer 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 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 the Weights section below for information on using weights with data from Topical Questionnaire sections.

The CPQ uses a three-month reference period; thus, the items administered in Summer 2019 (Round 84) asked beneficiaries about pain experienced in 2019. However, because the CPQ is administered to beneficiaries who were ever enrolled in Medicare in 2018 and are still enrolled in 2019, the CPQ data are released with the 2018 Survey File.

## **10.10 Demographics (DEMO)**

### *10.10.1 Core Content*

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

### *10.10.2 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 2015 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.<sup>26</sup>

The MCBS includes two ADI values for each beneficiary, one is the national-level percentile (ADINATNL) and the other is the state-level decile (ADISTATE). Both rankings are based on the census block group for the beneficiary's primary residence address (CENSBLCK). Beneficiaries have a value for each of these variables if

<sup>26</sup> University of Wisconsin School of Medicine and Public Health. 2015 Area Deprivation Index v2.0 Downloaded from <https://www.neighborhoodatlas.medicine.wisc.edu/>

their census block group is found on the ADI dataset. Excluding the exiting 2014 panel cases, there was a 93.6 percent match rate for cases matched to the ADI dataset.

**H\_DOB, H\_DOD, H\_AGE, and D\_STRAT:** The MCBS furnishes four variables related to the beneficiary's age in the DEMO segment. 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, 2018, 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 was entered (D\_DOB) in the MCBS instrument, the CAPI questionnaire automatically calculated the person's age, which was then verified with the respondent. In spite of 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 was asked again, in the next interview, to provide a date of birth. Some recording errors have been identified this way, but in most cases beneficiaries provided the same date of birth both times they were asked. In some cases, proxies indicated that no one was 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 was used to select and stratify the sample.

**D\_RACE2:** Race categories are self-reported by the respondent. Categories are not suggested by the interviewer, nor did 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 participant'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. RUCA codes are many, but permit stricter or looser delimitation of metropolitan, micropolitan, and small town commuting areas. 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.

### Primary and Secondary RUCA Codes

The 10 whole numbers shown in Exhibit 10.9.2a 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 was 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 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 10.10.2a: 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 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 10 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 (table 2). 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 10.10.2b:** Secondary RUCA (H\_RUCA2) Codes, 2010

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

HISPORIG: Hispanic/Latino origin includes persons of Mexican, Puerto Rican, Cuban, Central or South American or other Spanish culture or origin, regardless of race. These answers are self-reported by the respondent.

INCOME: Income represents the best source or estimate of income during 2018. Data gathered in fall and summer interviews represent the most detailed 2018 data and are used when available. For individuals not completing the Fall 2018 interview (that is, Continuing Panel people unavailable for Fall 2018), the most recent information available was 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. In some cases the respondent would not, or could not, provide specific information but did say the income was 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 was created following the rules below:

- Beneficiaries were assigned an INT\_TYPE if they completed or partially completed an interview in at least one round in 2018. 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 18 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 2018, and did not have any completed/partially completed questionnaire data for 2018. 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 2018 segments that include 2018 non-response adjustments: ACCSSMED, CHRNPAIN, FOODINST, INCASSET, MCREPLNQ, RXMED, 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, due to the fact that the data for these segments is collected in 2019, beneficiaries may have moved from a facility in 2018 to the community in 2019 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 2017 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 2018 data segment even if the beneficiary currently resides in a facility in 2018, 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\_IND: The income-to-poverty ratio (IPR) was added in 2015. 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, e.g., 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 those in “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 1 (beneficiary living alone or in a facility) or 2 (beneficiary living with a spouse only) as the Income and Asset information is collected only from the beneficiary and the beneficiary's spouse. Medicare beneficiaries have slightly different poverty level indices used for program eligibility. The IPR\_IND uses the Medicare poverty thresholds for calculation.

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

SPCHNLNM: Respondents were asked to report all living children, whether stepchildren, natural, or adopted children.

SPMARSTA: The respondent was allowed to define marital status categories; there was no requirement for respondents to report a legally recognized arrangement (e.g., married, divorced, etc.).

SPSDTH: Indicates if a respondent's spouse died within the last year. This variable was new in 2017.

SPVARATE: The VA disability rating variable 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>).

SURVIVE: This variable contains data from beneficiaries who were continuously enrolled in Medicare from January 1 up to and including their fall round interview.

### *10.10.3 Special Notes*

The DEMO segment now contains all demographic data from both the survey and from CMS administrative records. Beginning in 2016, an LEP variable indicates what language the respondent prefers to read in. H\_CBSA replaces H\_URBRUR beginning in 2017.

There are several changes to the ADI variables in 2018. These stem from changes to the input data source, which was updated prior to 2018 MCBS data processing. The ADI variables now contain national-level percentile and state-level decile rankings, rather than raw ADI values. Additionally, ADI scores are identified at the census block group only. Variables ZIPADI, AZIPCODE, and ZIPRSLT are no longer created. Variable CENSADI has been replaced by ADINATNL and ADISTATE.

## **10.11 Diabetes (DIABETES)**

### *10.11.1 Core Content*

This 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 CHRNCND 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.

### *10.11.2 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, yielding the numeric frequency (variables D\_INSFRQ, D\_INSPMP, D\_MEDFRQ, D\_TSTFRQ, and D\_SORFRQ, respectively). With the exception of D\_INSPMP, each variable has a corresponding frequency unit (variables INSUUNIT, MEDSUNIT, TESTUNIT, and SOREUNIT, respectively).

### *10.11.3 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, CHRNCND stores variables relevant to diabetes diagnoses (e.g., OCBETES and D\_OCDTYP). Variables related to diabetes risk and screening (e.g., DIAEVERT, DIARECNT, DIAAWARE, DIARISK, and DIASIGNS) appear in the PREVCARE segment. The variable pertaining to diabetic retinopathy (ERETINOP) appears in the VISHEAR segment.

## **10.12 Facility Assessments (FACASMNT)**

### *10.12.1 Core Content*

CMS designed the Minimum Data Set (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 nursing homes. Therefore, the MCBS is often able to abstract information applicable to the MCBS directly from the MDS. For this reason, the MCBS facility questionnaire has been designed to mirror the MDS instrument.

### *10.12.2 Variable Definitions*

#### **Activities of Daily Living (ADL) and Instrumental Activities of Daily Living (IADL):**

ADLs and IADLs: The MCBS asks whether respondents have any difficulty performing several activities. Their answers about difficulty performing the ADLs (PFBATHNG, PFDRSSNG, PFEATING, PFTRNSFR, PFLOCOMO, and PFTOILET) and IADLs (DIFUSEPH, DIFSHOP, DIFMONEY, IADSTOOP, IADLIFT, IADREACH, IADGRASP, and IADWALK) reflect whether or not the beneficiary usually had 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 were considered. If a respondent only performed an activity with help from another person (including just needing to have the other person present while performing the activity), or did not perform the activity at all, then that person was 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 were

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," were coded "yes."

**DIFMONEY:** 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.

**DIFSHOP:** Shopping for personal items means going to the store, selecting the items, and getting them home. Having someone accompany the respondent would qualify as help from another person.

**DIFUSEPH:** Using the telephone includes the overall complex behavior of obtaining a phone number, dialing the number, talking and listening, and answering the telephone.

**PFBATHNG:** Those 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 hand rails 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

**PFDRSSNG:** Dressing is the overall complex behavior of getting clothes from closets and drawers and then putting the clothes on. Tying shoelaces is not considered part of dressing, as is putting on socks or hose. 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.

**PFEATING:** 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.

**PFLOCOMO:** Walking means using one's legs for locomotion without the help of another person or special equipment or aids such as a cane, walker, or crutches. Leaning on another person, having someone stand nearby in case help is needed, and using walls or furniture for support all count as receiving help. Orthopedic shoes and braces are special equipment.

**PFTOILET:** Using the toilet is the overall complex behavior of going to the bathroom for bowel and bladder function, transferring on and off the toilet, cleaning after elimination, and arranging clothes. Elimination itself, and consequently incontinence, are not included in this activity, but were asked as a separate question, discussed next.

**PFTRNSFR:** Getting in and out of chairs includes getting into and out of wheelchairs. If the beneficiary holds onto walls or furniture for support, he or she is considered to receive "help from special equipment or aids," since the general population does not use such objects in getting in and out of chairs. Special equipment includes mechanical lift chairs and railings.

### 10.12.3 Special Notes

#### 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.

**Exhibit 10.12.1:** 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	
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

#### What is the difference between the MDS and FACASMNT data?

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

**Exhibit 10.12.2:** Differences between FACASMNT and MDS3 Data

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

**10.13 Facility Characteristics (FACCHAR)***10.13.1 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.

*10.13.2 Variable Definitions*

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

*10.13.3 Special Notes*

There were several variables that had the same variable names as items in NAGIDIS. These variables were renamed in 2018 to clearly indicate that the facility offers the assistance. The text below indicates the variable name used in 2017 and prior (left) and the new variable name in 2018 (right).

HELPBATH = BATHHELP  
 HELPDRES = DRESHHELP  
 HELPSHOP = SHOPHELP  
 HELPWALK = WALKHELP  
 HELPEAT = EATHHELP  
 HELPCOM = COMHELP

## 10.14 Falls (FALLS)

### *10.14.1 Core Content*

This file contains responses related to injuries and attitudes related to falls.

### *10.14.2 Variable Definitions*

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

### *10.14.3 Special Notes*

N/A

## 10.15 Food Insecurity (FOODINS)

### *10.15.1 Core Content*

This file contains information regarding the beneficiary's access to sufficient food. These questions are part of the Income and Assets Questionnaire (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>.

### *10.15.2 Variable Definitions*

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

### *10.15.3 Special Notes*

This questionnaire is administered the summer following the year of interest. The food insecurity section for the reference year 2018 was asked in the summer of 2019. 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 the Weights section below for information on using weights with data from Topical questionnaire sections.

## 10.16 General Health (GENHLTH)

### *10.16.1 Core Content*

This file contains data regarding a beneficiary's general health status and functioning such as height and weight.

### *10.16.2 Variable Definitions*

HEIGHTFT and HEIGHTIN: For height and weight, the respondent was 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.

HELMTACT: Limitations on activities and social life reflect the respondent's experience over the preceding month, even if that experience was atypical.

WEIGHT: 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.

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

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

Then categorized as:

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

### 10.16.3 Special Notes

N/A

## 10.17 Health Insurance Summary (HISUMRY)

### 10.17.1 Core Content

The Health Insurance Summary file contains information on administrative plans and their characteristics. Specifically, it includes flags for monthly enrollment and dual eligibility status, as well as 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.

### 10.17.2 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 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 dual eligible/Medicaid Only Dual 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: PDP Employer Group Waiver Plan Indicator: If the plan is an EGWP, then the value is 1, else the value will be 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 of the beneficiaries in the MCBS sample belong to Medicare managed care plans. CMS derived variables that describe this Medicare managed care membership (H\_GHPSW and 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 2018 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, as of the 2009 MCBS files. 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 the MARX data.

H\_MCSW: State buy-in is tracked by CMS and is 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. In 2015, CMS modified the data types of several variables. One example is the 2013 variable OP\_MDCD compared to the 2015 variable H\_OPMDCD. The variable values "1", "2", "3", "4" had been stored as character values. As part of the modifications applied in 2015, these values are stored as numeric values. The code definitions are equivalent. Users that want to merge the HISUMRY segment across Access to Care (2013 and prior) and Survey File (2015 and later) must change the data type of one variable and rename it.

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, -QDWI, -QI).

Medicaid Questions: To help the respondent answer the questions about Medicaid, the interviewers used the name of the Medicaid program in the state where the beneficiary was living. 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, you will need to recode the 2015 and later data to the previous values.

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

- 1 - No premium subsidy
- 2 - 25% premium subsidy
- 3 - 50% premium subsidy
- 4 - 75% premium subsidy
- 5 - 100% premium subsidy

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

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

To compare the 2015 and later H\_PDSL01-12 values, they will need to be mapped back to the 2013 and earlier values as listed in Exhibit 10.17.1.

**Exhibit 10.17.1:** Mapping of H\_PDLS01-12 Values from 2013 and Prior to 2015 and Beyond

2015 and later H_PDSL01-12 values*	Mapping to 2013 and earlier values
01,02,03,04,05	05
06	04
07	03
08	02
09,10,13	01

SOURCE: HISUMRY segment (2015 and beyond), RICA (2013 and prior)

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

H\_DDED01-12: The monthly values reflect the ANNUAL Part D deductible amount charged by the plan that the beneficiary was enrolled in on that month.

H\_PTAPRM: Total Part A Premium paid in 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\_PTBPRM: 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.

MTFCOVER: In 2017, this variable was dropped due to an instrument issue that resulted in some in universe beneficiaries not being asked the question. The variable will be re-introduced in future years.

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 example, 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. Prior to 2015, this information was released as TOT\_PREM. This new variable name highlights some of the uncertainty that may exist around the total premium paid for some plans. In addition, due to a change in the source of premium information for Part A, B, C, and D plans, data for 2015 and beyond are not directly comparable with data for earlier years.

## Payment Model Participation Flags

There are three variables that indicate the payment model for each plan. These payment model participation flags replace the 2015 and 2016 H\_ACOFLG variable that only indicated participation in the value '08' (Medicare Shared Savings Program or MSSP). The new variables show participation in all programs, not just MSSP.

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, 2<sup>nd</sup>, or 3<sup>rd</sup> 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.

### 10.17.3 Special Notes

When describing dual 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 (dual) 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)

■ **How do the 2015-2018 Health Insurance Timeline (HITLINE) and Health Insurance Summary (HISUMRY) segments differ from the previously released segments (i.e., RICs 4 and A)?**

► File Structure:

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

- d. Part D premiums: This is the dollar amount of the Part D Total Premium (basic + supplemental) Rate (Net of Rebates). The Part D Total Premium is the sum of the Basic and Supplemental Premiums. This amount is net of any Part A/B rebates applied to "buy down" the drug premium for Medicare Advantage plans; for some plans, the total premium may be lower than the sum of the basic and supplemental premiums due to negative basic or supplemental premiums. The data source is constructed from information submitted by Part D plan sponsors to CMS's Health Plan Management System (HPMS). Employer and National Program of All-inclusive Care for the Elderly (PACE) plans are waived from reporting Plan Benefit Package information. For those plans that did not report, the value of this variable will be blank.
  - e. Survey Only Plan premiums: This information is sourced from the D\_ANNPRM variable in the HITLINE file. D\_ANNPRM is the annualized premium that the beneficiary pays for their health insurance plan. It is obtained from the survey and reflects information for survey-reported plans ONLY. D\_ANNPRM is set to missing for all administrative plans. Premium information associated with a plan would not be included in D\_ANNPRM (and, hence, EST\_TPRM) unless a beneficiary reported both the premium amount paid and the frequency of those payments. Data users interested in making different assumptions about how premium information for survey-reported plans should be handled can utilize information provided in the HITLINE segment, as it contains all available data for each survey-reported plan. Note that because MCBS Facility interviews are conducted with facility staff rather than beneficiaries, the survey has no information on premiums paid by these beneficiaries for survey-only plans. For beneficiaries living in a facility setting, EST\_TPRM reflects Part A – D premiums only.
- Private plans reported in a community setting can be linked across years using the unique plan identifier, PLANNUM.
  - HITLINE files now indicate why certain information is missing for a given plan. The value of 'N' (Not ascertained) refers to instances where a beneficiary was supposed to receive a question but did not, which can occur for several reasons. The value of '.' (Inapplicable/Missing) refers to instances where a beneficiary was out of universe for a given question (for example, all beneficiaries who report not paying a premium will have a value of '.' for the premium amount associated with that plan).
- ▶ Plans Reported for Beneficiaries Living in a Facility:
    - Plans reported for beneficiaries living in a facility setting have fewer details about coverage collected, which is a consequence of data being collected from facility staff, rather than the beneficiary. The HITLINE files include a flag, D\_FCLTYF, which makes it possible to identify plans that are missing certain plan details because they were reported in a facility setting.
    - Given that only limited information about coverage type is collected from facility staff, for beneficiaries who cross-over from the community to a facility or move from one facility to another, it is not possible to determine whether a plan reported in a facility is, in fact, the same plan that was also reported in a community setting and / or a plan reported in a facility setting at a different point in time.
  - ▶ File-Specific Notes:

- PLANTYPE can change from year to year for some plans. This is a consequence of beneficiaries (or a proxy) providing slightly different information about a plan, which can result in changing classification from year to year. While responses are broadly consistent over time regarding whether a plan provides comprehensive coverage, there is less consistency regarding whether the plan was employment-based or purchased directly.

## 10.18 Health Insurance Timeline (HITLINE)

### 10.18.1 Core Content

This 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. In addition, HITLINE also 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 through the Department of Veteran's Affairs, 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.

Prior to 2013, detailed information on survey-reported plans was available only for the first five plans and was included in the HISUMRY segment. In the 2015 Survey File LDS and beyond, this content is available for all plans and will appear in HITLINE.

### 10.18.2 Variable Definitions

PLANNUM: Unique plan number associated with a plan reported in a Community setting.

PLANTYPE: Indicates the type of plan.

S\_INS: Specifies whether the private health insurance plan has limited service coverage, such as dental-only, prescription drug-only, etc. This information was developed through an editing process in which plan names were researched and combined with other survey-reported plan information.

BEGDATE: The date the plan coverage began.

ENDDATE: The date the plan coverage ended

SCRCOV01-12: Indicates the source of coverage information for the plan: 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\_HMOPPO: Indicates whether beneficiary's private plan is an HMO/PPO. Obtained from (HI25) or (HIS25).

S\_PHREL: The relationship of the policyholder to the beneficiary. Responses from (HIS26) or (HI26) 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 either (HI27) or (HIS27).

S\_COVNM: The number of people covered by each private plan. This information is obtained from either (HI29) or (HIS29).

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 (HI31A) or (HIS31A).

S\_IP: Indicates whether beneficiary's private plan covers inpatient stays. Obtained from (HI31A) or (HIS31A).

S\_COVNH: Indicates whether beneficiary's private plan has long-term care coverage. This information is obtained from either (HI31A) or (HIS31).

S\_DNTAL: Indicates whether beneficiary's private plan covers dental services. Obtained from (HI31A) or (HIS31A).

S\_PAYSP: Does the main insured person (MIP) pay any part of the insurance premium? Obtained from either (HI32) or (HIS32).

S\_PREM: Reported cost of private health insurance plan premiums. A premium amount was recorded even if the respondent did not directly pay the premium (if, for example, a son or daughter paid the premium). This variable was derived from responses to (HI33) and (HIS33). For family plans, the reported amount reflects the total premium paid for the plan.

D\_ANNPREM: The annual reported cost of private health insurance plan premiums calculated for beneficiaries who answered questions associated with both S\_PAYUNIT and S\_PREM. Premium amounts have been prorated based on how long the beneficiary held the policy. For family plans, the annualized amount reflects the total premium paid for the plan.

S\_PAYUNT: Specifies how frequently (once per year, once per month, etc.) the amount reported in S\_PAYSP was paid. This information is based on either (HI33) or (HIS33).

S\_PAYOTH: Indicates whether anyone else, such as an employer or a union, helped to pay any portion of the premium. Obtained from either (HI33A) or (HIS33A).

S\_PAYWHO: Indicates who paid a portion of the total cost of the premium. Obtained from (HI33B) and (HIS33B).

S\_TRIRX: Specifies where Tricare members obtain prescription drugs. This information is derived from either (HIT4A1) or (HIST3AA).

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.

### *10.18.3 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 (If someone had coverage Jan – March and June – November, BEGDATE will reflect that coverage started in January). Most plans are going to have an ENDDATE, as well. The only plans with missing ENDDATE will be 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.

The variables D\_PRIVAT, D\_HMO\_COV, D\_HMOCUR, D\_MCAID, D\_MCARE, and D\_MCRHMO from previous releases of the MCBS are not on this file; however, the information regarding Medicare Advantage coverage, Medicaid coverage, Medicare coverage, and the sources of the coverage information is contained within the file.

Starting with 2015, the HITLINE segment is provided in a long format versus a wide format. If data users are conducting analysis covering a period that includes both pre-2015 and 2015 or later data years, data should be transposed in order to merge them. See the codebook for explanations of codes for plan types and details about the private plans reported.

For more information on how the 2015-2018 HITLINE and HISUMRY segments differ from the previously released segments (i.e., RICs 4 and A), please see section 10.16.3 above.

## 10.19 Household Characteristics (HHCHAR)

### 10.19.1 Core Content

This file 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, the file reflects the latest available data on the size of the household and the age and relationship of the people in it. Information about the beneficiary's physical residence is collected at the Baseline interview and updated as necessary.

### 10.19.2 Variable Definitions

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 or exact relationship of a household member was sometimes unknown (perhaps because a proxy provided the information), the sum of the variables "number related"/"number not related" (D\_HHREL/D\_HHUNRL) or "number under 50"/"number 50 or older" (D\_HHLT50/D\_HHGE50) may not equal the total number of people in the household (D\_HHTOT).

Starting with 2017, HHCHAR makes available expanded definitions of household composition (D\_COMPHH) and sex of spouse/partner (D\_SEXSPP). Variable D\_COMPHH "Household Composition" replaces the

variable D\_HHCOMP. The new variable still reflects the composition of household members but contains a much broader array of relationships. The variable D\_SEXSPP "Sex of Spouse/Partner" indicates the sex when a spouse or partner is identified in D\_COMPHH as a member of the household.

### 10.19.3 Special Notes

As of 2017, information about whether the beneficiary has a residence in another state is no longer included in the file. 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.

A new variable was added to HHCHAR for the 2018 data year: SPMOVED "Has the SP moved since the last Fall Round data collection date?" Only beneficiaries living in the community who are responding to a *Continuing* interview are in universe for this question. 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.

## 10.20 Income and Assets (INCASSET)

### 10.20.1 Core Content

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

### 10.20.2 Variable Definitions

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

### 10.20.3 Special Notes

In the IAQ, the reference period for income is generally the previous calendar year. That is, many income questions are asked in the summer of 2019 about income earned in 2018.

- Example: "Now I want to ask about your [and spouse'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 the Weights section below for information on using weights with data from Topical questionnaire sections.

The MCBS imputes income in 2018 when income data are missing. Data were first imputed whether or not an income source (such as Social Security) existed. If the income source exists, then the amount earned was imputed next. Imputation was performed using the hot deck imputation method, and a flag was created for each imputed variable indicating whether or not the corresponding value was imputed.

## 10.21 Interview Characteristics (INTERV)

### 10.21.1 Core Content

This segment summarizes the characteristics of the interview such as the type of interview conducted and whether or not a proxy was used.

### 10.21.2 Variable Definitions

**INTERVU:** There is one record for each individual for each round of completed interviews, either in the community (INTERVU = "C") or a facility (INTERVU = "F"). Some beneficiaries had more than one interview in a round. 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 was conducted.

**INTVDISP:** The respondent's status as of the last day of the calendar year. There are two dispositions (40 = complete and 50 = complete-deceased).

### Proxy Rules:

Proxy respondents were always used in nursing homes, homes for beneficiaries with intellectual disabilities, and psychiatric hospitals. The need for a proxy when interviewing respondents in other institutions was evaluated on a case-by-case basis.

In long-term care facilities, the proxy respondents were members of the staff at the facility identified by the administrator. Usually, more than one respondent was used; for example, a nurse may have answered the questions about health status and functioning, while someone in the business office handled questions about financial arrangements.

**SPPROXY:** People who were too ill or who could not complete the Community interview for other reasons were asked to designate a proxy. A proxy is someone very knowledgeable about the beneficiary's health and living habits. In many cases, the proxy was a close relative such as the spouse, a son, or daughter. In other cases, the proxy was a non-relative like a close friend or caregiver.

Proxy variables:

- ROSTREL: Indicates the proxy's relationship to the beneficiary (e.g., spouse, son, daughter).
- SPPROXY: Indicates whether or not a Community interview was conducted with a proxy respondent or with the beneficiary.
- WHYPROXY: Indicates the reason that a proxy was needed.

### Other variables:

INTVDATE: Date on which the interview was conducted.

MINTOTAL: MINTOTAL contains the length of the interview, in minutes. A new timestamp was implemented in 2016 that captures actual questionnaire administration time only while excluding any time that elapses for setting up the interview.

TOTLINTV: Indicates the total number of interviews conducted with this beneficiary. Community interviews are sometimes interrupted to accommodate the respondent's schedule or for other reasons. Facility interviews are conducted with several instruments and often involve many respondents.

INTVFLG: Added in 2018, this flag indicates if an interview was completed in the fall.

### *10.21.3 Special Notes*

N/A

## 10.22 Medicare Advantage Plan Questions (MAPLANQX)

### *10.22.1 Core Content*

The Medicare Advantage (MA) Plan Questions segment augments information from the Access to Care (ACQ) and Satisfaction with Care (SCQ) sections of the questionnaire for those beneficiaries enrolled in Medicare Part C. Beneficiaries who are enrolled in a Medicare Advantage plan at the time of the interview are asked general questions about their health plans, which include access to and satisfaction with medical services. The file 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.

### *10.22.2 Variable Definitions*

D\_ANHMO: What is the annual additional cost of Medicare Advantage premiums? The premiums have been annualized regardless of the length of time the respondent actively held the policy. This variable is derived from the Health Insurance (HIQ) item MHMOAMT.

### *10.22.3 Special Notes*

N/A

## 10.23 Medicare Plan Beneficiary Knowledge (MCREPLNQ)

### *10.23.1 Core Content*

This segment contains information about the beneficiary's knowledge with the Medicare open enrollment period and knowledge about Medicare-covered expenses.

The data collected in this segment will allow an evaluation of the impact of existing education initiatives by CMS. The KNQ questionnaire section helps to 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.

### *10.23.2 Variable Definitions*

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

### *10.23.3 Special Notes*

This questionnaire is administered the winter following the year of interest. The KNQ questions for the reference year 2018 were asked in the winter of 2019. 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 the Weights section below for information on using weights with data from Topical questionnaire sections.

## **10.24 Mental Health (MENTHLTH)**

### *10.24.1 Core Content*

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

### *10.24.2 Variable Definitions*

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).

### *10.24.3 Special Notes*

N/A

## **10.25 Mobility (MOBILITY)**

### *10.25.1 Core Content*

This segment includes beneficiaries who resided in a Community setting at any point during the calendar year and contains information on the beneficiary's use of available transportation options and whether the beneficiary's health affects their daily travel.

### *10.25.2 Variable Definitions*

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

### 10.25.3 Special Notes

All variables in this section are derived by summarizing the data collected in each data collection round to produce information for the benefit year, and therefore all begin with the letter D. Beneficiaries who previously reported giving up driving are not asked about whether they have given up driving again. Instead, their information is carried forward from the previous year.

## 10.26 Minimum Data Set (MDS3)

### 10.26.1 Core Content

The Minimum Data Set is assessment information collected while the beneficiary was in an approved Medicare Facility.

CMS designed the Minimum Data Set (MDS3) instrument to collect information regarding the health status and functional capabilities of nursing home residents. The MDS is administered to anyone residing in a certified nursing home, regardless of payer. For this reason, the MCBS Facility Instrument has been designed to mirror the MDS instrument. By adapting the applicable MCBS questions, interviewers can extract data directly from these assessments which expedites collection, while ensuring quality.

For more information regarding the MDS and the changes in version 3.0, please consult

<https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/NursingHomeQualityInits/index>.

### 10.26.2 Variable Definitions

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

### 10.26.3 Special Notes

MDS3 records are included for a beneficiary having such a record in the year of interest.

What is the difference between the MDS and FACASMNT data? See [Facility Assessments](#) (FACASMNT) section above.

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.

## 10.27 NAGI Disability (NAGIDIS)

### 10.27.1 Core Content

This 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.

### 10.27.2 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 had 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 were 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 was 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 were 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," were 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: Those 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 hand rails 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 is not considered part of dressing as is putting on socks or hose. 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) was made clear by examples. Washing dishes, straightening up and light cleaning represent light housework; scrubbing floors and washing windows represent heavy housework. The interviewer was 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.

### *10.27.3 Special Notes*

Since 2016, six global disability questions are released to comply with HHS guidance. The variables are: DISDECSN, DISWALK, DISBATH, and DISERRND. 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.

## **10.28 Nicotine and Alcohol (NICOALCO)**

### *10.28.1 Core Content*

The Nicotine and Alcohol questionnaire (NAQ) section was first added to the MCBS in 2016 to follow HHS guidelines for asking about alcohol and nicotine consumption. This segment contains NAQ information on the prevalence and frequency of alcohol and nicotine use (including cigarettes, e-cigarettes, cigars, pipe tobacco, and smokeless tobacco).

### *10.28.2 Variable Definitions*

Please see the Codebook for variables included in this segment.

### *10.28.3 Special Notes*

N/A

## 10.29 Outcome and Assessment Information (OASIS)

### 10.29.1 Core Content

This 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>.

### 10.29.2 Variable Definitions

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

### 10.29.3 Special Notes

All home health records are included for MCBS participants for the year of interest.

## 10.30 Patient Activation (PNTACT)

### 10.30.1 Core Content

The data in this segment 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 not only if beneficiaries receive information about their health and Medicare, but also if they understand it in a way that makes it useful.

### 10.30.2 Variable Definitions

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

### 10.30.3 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.

## 10.31 Preventive Care (PREVCARE)

### 10.31.1 Core Content

This 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, and shingles vaccine.

### 10.31.2 Variable Definitions

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

### 10.31.3 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.

## 10.32 Residence Timeline (RESTMLN)

### 10.32.1 Core Content

With the 2018 data year, the RESTMLN segment has been slightly modified. The RESTMLN segment provides a timeline of each MCBS setting type in which a beneficiary resided over the portion of the year in which they were enrolled in Medicare, as well as any periods associated with FFS inpatient, SNF, or Hospice events. The total number of setting changes is equal to the sum of MCBS residential status changes (D\_NUMSIT) 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. Residential status situations will not have end dates populated to illustrate that these extend through any claim events which follow until a change in residential status occurs.

### 10.32.2 Variable Definitions

Please see the Codebook for information regarding variables in this segment. There are some variables on this segment that were derived or created by combining two or more survey variables. These derived variables are preceded with the characters "D\_". Variables are created or modified in order to recode data items, to protect the confidentiality of survey participants, or to globally edit some values.

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 file year. As mentioned above, this is calculated by adding D\_NUMSIT and D\_NUMEVT. At a minimum, each beneficiary will have information pertaining to their setting at the beginning of their eligibility period within the file year, as described below.

The variable D\_BEG1 will represent the beneficiary's first date of Medicare eligibility within the file year. The associated D\_CODE1 variable will either identify a residential setting or for a small number of cases, contain the code 'N.' The latter only occurs for some facility respondents who are new participants in the MCBS survey, but were enrolled in Medicare prior to the start of the file 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 1<sup>st</sup> of the file year, it will result in the setting code on their first situation (D\_SIT1) having a value of 'N.'

### 10.32.3 Special Notes

The 2015 segment was included with the Cost Supplement LDS for data year 2015 only, as it was constructed using the Cost Supplement weights. For the 2016 data year and beyond, the segment is included with the Survey File LDS.

In 2018, a couple of changes were made to how the timeline is constructed. Previously, residential situations were overwritten by all claim events which overlapped them, regardless of claim type or when the overlap occurred. Beginning with this file year, Hospice events only will no longer 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. They should instead be considered as occurring concurrently with the beneficiary's identified residential situation. Also, a beneficiary's initial residential status will not be 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.

## 10.33 RX Medications (RXMED)

### 10.33.1 Core Content

The RXMED segment augments information from the Access to Care (ACQ) and Satisfaction with Care (SCQ) sections of the questionnaire with information specific to prescription drug coverage. Items asked in the Drug Coverage Questionnaire (RXQ) cover 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.

### 10.33.2 Variable Definitions

**OPTIONS CONSIDERED WHEN CHOOSING PRESCRIPTION DRUG COVERAGE:** These seven items were asked as a series of separate yes/no questions: PDOPTDUC, PDOPTFOR, PDOPTGAP, PDOPTPAY, PDOPTPRE, PDOPTREC, and PDOPTVEN. If the respondent answers “yes” to any of these items, the respondent was then asked at PDPMOST to select which of the seven was the most important consideration when they thought about options for drug coverage.

**REASONS DID NOT USE CURRENT COVERAGE:** These six items were asked as a single question (RXQ item PDNOUSE), with multiple responses allowed. This question is asked of respondents who indicated that they did not use their current prescription drug coverage (variable RXUSEPLN).

**REASONS NOT ENROLLED IN MEDICARE PRESCRIPTION DRUG PLAN:** These 12 items are derived from a single question (RXQ item PDNTENR), with multiple responses allowed. This question is asked for beneficiaries who do not have any of the following: a current Medicare prescription drug plan, a current Medicare managed care plan that has prescription coverage, or a current private plan that has prescription coverage.

### 10.33.3 Special Notes

This questionnaire is administered the summer following the year of interest. The RXQ questions for the reference year 2018 were asked in the summer of 2019. 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 the Weights section below for information on using weights with data from Topical questionnaire sections.

## 10.34 Satisfaction with Care (SATWCARE)

### 10.34.1 Core Content

This segment contains data on satisfaction with health care and reasons why beneficiaries do not seek medical care or purchase prescription drugs.

### 10.34.2 Variable Definitions

The questions about satisfaction with care represent the respondent's general opinion of all medical care received in the year preceding the interview.

**MCDRNSEE:** If a respondent mentioned any health problem that was not cared for, it was recorded without discrimination; the respondent might have referred to a small ache or pain, or to a serious illness or symptom.

**REASONS FOR NOT OBTAINING PRESCRIPTION:** These items are asked for respondents who indicated in item PMNOTGET that there were prescriptions that they did not obtain. A single select-all-that-apply question (SCPMCODS) is initially asked, with ten possible response options and multiple responses allowed, which are then coded into the ten relevant Yes/No variables in the SATWCARE segment: SCPMCOST, SCNOHELP, SCPMREAC, SCPMNLKE, SCPMNCND, SCPMNOCV, SCPMTROB, SCPMSMPL, SCPMSUBS, and SCPMOTHR. In addition, the following question (SCPMMAIN) asks the main reason why the prescription was not obtained, with one response allowed.

### *10.34.3 Special Notes*

Verbatim questions MCDISVB, SCROTOS, and SCPMOTOS were back coded into response categories as necessary; verbatim text is not released.

## **10.35 Usual Source of Care (USCARE)**

### *10.35.1 Core Content*

This segment contains data on where and how the beneficiary typically seeks medical care.

### *10.35.2 Variable Definitions*

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

### *10.35.3 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 the Weights section below for information on using weights with data from Topical questionnaire sections.

## **10.36 Vision and Hearing (VISHEAR)**

### *10.36.1 Core Content*

This file contains information on the beneficiary's eye health and hearing status.

### *10.36.2 Variable Definitions*

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

### *10.36.3 Special Notes*

Since 2016, 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.

## 10.37 Weights

### 10.37.1 Cross Sectional Weights

Two types of weights are provided, cross-sectional weights and longitudinal weights. Cross-sectional weights apply to the entire file of all those who completed an interview, either Community or Facility. The first set of cross-sectional weights (CENWGTS) can be used for making estimates of the population of Medicare beneficiaries who were continuously enrolled in Medicare from January 1<sup>st</sup> up to and including their fall interview (i.e., the “continuously enrolled” population). The second set of cross-sectional weights (EVRWGTS) can be used for making estimates of the population of Medicare beneficiaries who were enrolled in Medicare at any time during the entire calendar year (i.e., the “ever enrolled” population). Cross sectional weights are available for the Survey File and the Cost Supplement File in each data year. The Survey File LDS contains weights for the continuously enrolled (CENWGTS) and ever enrolled (EVRWGTS) populations.

### 10.37.2 Longitudinal Weights<sup>27</sup>

Longitudinal weights allow for the study of respondents across data years.

Two-year longitudinal weights (LNG2WGTS) 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 apply to members of the 2015, 2016, and 2017 panels who were alive and entitled as of the Fall 2018 (Round 82) interview, had 2017 and 2018 Survey File data, and enrolled on or before 1/1/2017. 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.

Three-year longitudinal weights (LNG3WGTS) 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 for members of the 2015 and 2016 panels who were alive and entitled as of the Fall 2018 (Round 82) interview, had Survey File data in 2016 and 2018, enrolled on or before 1/1/2016, and were continuously enrolled through the fall of 2018 (i.e., three years). 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.

Four-year longitudinal weights (LNG4WGTS) 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 for members of the 2015 panels who were alive and entitled as of the Fall 2018 (Round 82) interview, had Survey File data in 2015 and 2018, enrolled on or before 1/1/2015, and were continuously enrolled through the fall of 2018 (i.e., three years). 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.

For a further discussion about the ever enrolled and continuously enrolled populations 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>.

<sup>27</sup> 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.

### *10.37.3 Using weights with data from Topical Questionnaire Sections*

To generate estimates using the data from one of the eight Topical Questionnaire sections (FOODINS, INCASSET, MCREPLNQ, RXMED, ACCSSMED, USCARE, PNTACT, and CHRNPAIN) on their own or merged with another Survey File segment that does not contain special non-response adjustment weights, the analyst must always use the special non-response adjustment general and replicate weights included in the Topical segment **INSTEAD** of using the general and replicate weights that appear in the separate weight segments (CENWGTS, EVRWGTS). 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 sections. Each segment with data from Topical questionnaire sections 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 Survey File ever 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 section the number of beneficiaries with a continuously enrolled Topical weight is larger than the number of beneficiaries with an ever enrolled Topical weight.

## 11. REFERENCES

- Eicheldinger, Celia, and Arthur Bonito. "More accurate racial and ethnic codes for Medicare administrative data." *Health care financing review* 29, no. 3 (2008): 27.
- Kalton, Graham, and Daniel Kasprzyk. "The Treatment of Missing Survey Data." *Survey methodology* 12, no. 1 (1986): 1-16.
- University of Wisconsin School of Medicine and Public Health. *2015 Area Deprivation Index v2.0* Downloaded from <https://www.neighborhoodatlas.medicine.wisc.edu>
- Wolter, Kirk. *Introduction to variance estimation*. Springer Science & Business Media, 2007.

# APPENDICES

## 12. APPENDICES

### Appendix A: MCBS Common Definitions

**Activities of daily living (ADLs):** Activities of daily living are activities related to personal care. They include bathing or showering, dressing, getting in and out of bed or a chair, walking, using the toilet, and eating.

**Baseline interview:** The initial questionnaire administered to new respondents to the study; administered in the fall of the year they are selected into the sample (interview #1).

**Beneficiary:** Beneficiary refers to a person receiving Medicare services who may or may not be participating in the MCBS. Beneficiary may also refer to an individual selected from the MCBS sample about whom the MCBS collects information. Beneficiaries must meet at least one of three criteria for Medicare eligibility (is aged 65 years or older, is under age 65 with certain disabilities, or is of any age with End-Stage Renal Disease) and is entitled to health insurance benefits. (Source: <https://www.cms.gov/Medicare/Medicare-General-Information/MedicareGenInfo/index.html>).

**Claim-only event:** A claim-only event is a medical service or event known only through the presence of a Medicare Fee-for-Service claim from administrative data. This means that the event represented in the data could not be reconciled with a corresponding survey-reported event.

**Community component:** Survey of 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.

**Company clinic:** A doctor's office or clinic, which is operated principally for the employees (and sometimes their dependents) of a particular company or business.

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

**Continuously enrolled (aka always enrolled):** A Medicare beneficiary who was enrolled in Medicare from the first day of the calendar year until the fall interview and did not die prior to the fall round. This population excludes beneficiaries who enrolled during the calendar year 2015, those who dis-enrolled or died prior to their fall interview, residents of foreign countries, and residents of U.S. possessions and territories.

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

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

**Current-year enrollee:** Beneficiaries who were eligible and enrolled in Medicare (Parts A or B) anytime from January 1 to December 31 of the year the sample was selected.

**Doctor:** This includes both medical doctors (M.D.) and doctors of osteopathy (D.O.). It does not include chiropractors, nurses, technicians, optometrists, podiatrists, physician's assistants, physical therapists, psychologists, mental health counselors or social workers. Generic specialties shown in parentheses following one of the specialties were coded as the specialty. For example, if the respondent mentioned a "heart" doctor, cardiology was coded. Generic answers not listed were not converted to specialties.

**Doctor's office or group practice:** This refers to an office maintained by a doctor or a group of doctors practicing together; generally the patient makes an appointment to see a particular physician.

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

**Exit interview:** Conducted in the winter round, this interview completes the respondent's participation in the MCBS (interview #11). The exit interview is a special case of the Continuing interview.

**Facility component:** Survey of respondents 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.

**Fee-for-Service (FFS) payment:** Fee-for-Service is a method of paying for medical services in which each service delivered by a provider bears a charge. This charge is paid by the patient receiving the service or by an insurer on behalf of the patient.

**Field interviewer:** The principal contact for collecting and securing respondent data.

**Field manager:** A supervisor who motivates and manages a group of field interviewers to meet the goals of high quality data collection on time and within budget limits.

**Free-standing surgical center:** A facility performing minor surgical procedures on an outpatient basis, and not physically connected to a hospital. Note that a unit performing outpatient procedures connected with a hospital (either physically or by name) is referred to as a hospital outpatient department/clinic.

**Gap days:** Gap days are periods during the calendar year in which a sample person was enrolled in Medicare but was not covered by a survey interview.

**Home:** This includes situations where the doctor comes to the beneficiary, rather than the beneficiary going to the doctor. Here, "home" refers to anywhere the beneficiary was usually staying at the time of the medical provider's visit. It may be his/her home, the home of a friend, a hotel room, etc.

**Hospital emergency room:** This means the emergency room of a hospital. "Urgent care" centers are not included. (NOTE: All hospital emergency room visits were included, even if the respondent went there for a "non-emergency" condition such as a cold, flu or intestinal disorder.) A physician, nurse, paramedic, physician extender, or other medical provider may administer the health care.

**Hospital outpatient department:** A unit of a hospital, or a facility connected with a hospital, providing health and medical services, health education, health maintenance, preventive services, diagnosis, treatment, surgery, and rehabilitation to individuals who receive services from the hospital but do not require hospitalization or institutionalization. Outpatient clinics can include obesity clinics; eye, ear, nose and throat clinics; alcohol and drug abuse clinics; physical therapy clinics; kidney dialysis clinics, and radiation therapy clinics. The outpatient department may or may not be physically attached to a hospital, but it must be associated with a hospital.

**Incoming Panel Sample (formerly known as Supplemental Panel):** A statistically sampled group of beneficiaries that enter the MCBS in the fall of a data collection year. One panel is retired at the conclusion of each winter round, and a new panel is selected to replace it each fall round. Panels are identified by the data collection year (e.g., 2015 panel) in which they were selected.

**Instrumental activities of daily living (IADLs):** Instrumental activities of daily living are activities related to independent living. They include preparing meals, managing money, shopping for groceries or personal items, performing light or heavy housework, and using a telephone. If a beneficiary had any difficulty performing an activity by himself/herself, or did not perform the activity at all, because of health problems, the person was deemed to have a limitation in that activity. The limitation may have been temporary or chronic at the time of the survey. Facility interviewers did not ask about the beneficiary's ability to prepare meals or perform light or heavy housework, since they are not applicable to the beneficiary's situation; however, interviewers did question proxies about the beneficiary's ability to manage money, shop for groceries or personal items, or use a telephone.

**Internal Sample Control File:** A data file that contains every beneficiary sampled back through the beginning of MCBS. The file contains sampling information, year of selection, primary sampling unit, secondary sampling unit, contact information, and other sampling demographic information as well as final disposition codes to indicate completion status per round, component fielded per round, dates of death, and lost entitlement information.

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

**Medical clinic:** This refers to any group of doctors or other health professionals who have organized their practice in a clinic setting and work cooperatively; generally, patients either come in without an appointment or make an appointment and see whatever health professional is available.

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

- Hospital Insurance (Part A): covers inpatient hospital stays, care in a skilled nursing facility, hospice care, and some home health care.
- Medical Insurance (Part B): covers certain doctors' services, outpatient care, medical supplies, and preventive services.
- Medicare Advantage (Part C): an alternative to coverage under traditional Medicare (Parts A and B), a health plan option similar to a Health Maintenance Organization (HMO) or Preferred Provider Organization (PPO) administered by private companies.
- Prescription Drug Coverage (Part D): additional, optional coverage for prescription drugs administered by private companies.

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

**Medicare Advantage (MA):** Medicare Advantage Plans, sometimes called "Part C" or "MA Plans," are offered by private companies approved by Medicare. An MA provides, or arranges for the provision of, a comprehensive package of health care services to enrolled persons for a fixed capitation payment. The term "Medicare Advantage" includes all types of MAs that contract with Medicare, encompassing risk MAs, cost MAs, and health care prepayment plans (HCPs).

**Medicare beneficiary:** See Beneficiary.

**Medicare Managed Care Organization (MCO)/Health Maintenance Organization (HMO):** This is an organization that provides a full range of health care coverage in exchange for a fixed fee/co-pay. Some managed care plans require that plan members receive all medical services from one central building or

location. Formerly referenced only as HMOs, these organizations are now referred to with terms such as Medicare MCOs/HMOs/Medicare Advantage/Part C.

**Minimum Data Set (MDS):** The Minimum Data Set (MDS) is part of the federally mandated process for clinical assessment of all residents in Medicare and Medicaid certified nursing homes. For more information, please visit <https://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/Minimum-Data-Set-3-0-Public-Reports/index>.

**Neighborhood/family health center:** A non-hospital facility which provides diagnostic and treatment services, frequently maintained by government agencies or private organizations.

**Other clinic:** A non-hospital facility clinic that is not already listed in the other clinic categories. Some examples include a "free" clinic, a family planning clinic, or military base clinic.

**Outcome and Assessment Information Set (OASIS):** The instrument/data collection tool used by CMS to collect and report performance data by Medicare-certified home health agencies. For more information, please visit <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HomeHealthQualityInits>.

**Panel:** See Incoming Panel sample.

**Personal health care expenditures:** Personal health care expenditures consist of health care goods and services purchased directly by individuals. They exclude public program administration costs, the net cost of private health insurance, research by nonprofit groups and government entities, and the value of new construction put in place for hospitals and nursing homes.

**Prescription drugs:** The basic unit measuring use of prescription drugs is a single purchase of a single drug in a single container. Prescription drug data are included for beneficiaries living in the community; Prescription drugs administered during an inpatient hospital stay or to beneficiaries living in a facility are not included.

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

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

**Race/ethnicity:** Responses to race and ethnicity questions are self-reported by the respondent. Respondents who reported they were white and not of Hispanic origin were coded as white non-Hispanic; those who reported they were black/African-American and not of Hispanic origin were coded as black non-Hispanic; persons who reported they were Hispanic, Latino/Latina, or of Spanish origin, regardless of their race, were coded as Hispanic; persons who reported they were American Indian or Alaska Native, Asian, Native Hawaiian or other Pacific Islander, or two or more races and not of Hispanic origin were coded as other race/ethnicity.

**Reference Period:** The timeframe to which a questionnaire item refers.

**Residence status:** Medicare beneficiaries who only completed Community interviews during the calendar year are categorized as residing only in the community. Medicare beneficiaries for whom only Facility interviews were completed during the calendar year are categorized as residing only in facilities. Beneficiaries who completed at least one Community interview and for whom at least one Facility interview was conducted during the year are classified as residing in both community and facility.

**Respondent:** 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.

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

**Rural health clinic:** A clinic that provides outpatient services, routine diagnostic services for individuals residing in an area that is not urbanized and is designated as a health staff shortage area or an area with a shortage of personal health services. The clinic can also provide outpatient services that include physician services, services and supplies provided under the direction and guidance of a physician by nurse practitioner, physician assistants, and treatment of emergency cases. These services are usually provided at no charge except for the amount of any deductible or coinsurance amount.

**Sample person:** An individual beneficiary selected from MCBS' Incoming Panel sample to participate in the MCBS survey.

**Survey-reported event:** A survey-reported event is a medical service or event reported by a respondent during an interview. The event may have been matched to a Medicare Fee-for-Service claim from administrative data, or it may be a survey-only event, in which case it was not matched to a Medicare claim and is only known through the survey.

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

**Skilled nursing facility (SNF):** A facility (which meets specific regulatory certification requirements) which primarily provides inpatient skilled nursing care and related services to patients who require medical, nursing, or rehabilitative services but does not provide the level of care or treatment available in a hospital. (Source: <https://www.cms.gov/apps/glossary/default.asp?Letter=S&Language=English>)

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

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

**Walk-in urgent center:** A facility not affiliated with a nearby hospital, offering services for acute conditions (e.g., flu, virus, sprain). Typically, people are seen without appointments (i.e., walk-ins).

## Appendix B: 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

## Appendix C: Sample Code

### *Merging Segments within the 2018 Survey File LDS*

Data users can merge 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.

Below, is an example of how multiple Survey File segments can be merged with the CENWGTS segment in SAS using BASEID as the key variable. When merging segments, all observations in the CENWGTS segment should be preserved.

```
data merged;
  merge survey18.CENWGTS (in = a)
        survey18.DEMO (keep = BASEID H_AGE INT_TYPE)
        survey18.CHRNCOND (keep = BASEID D_OCCTYP)
        survey18.GENHLTH (keep = BASEID GENHELTH);
  by BASEID;
  if a;
run;
```

In order to segment the file to beneficiaries living in the community only, subset the file on the variable INT\_TYPE.

```
data merged_surveyfile;
  set merged;
  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 dual eligibles (i.e., enrolled in both Medicare and Medicaid programs) during CY2017 and CY2018.

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.<sup>28</sup> Therefore, the examples presented in this section involving multiple years of MCBS data use replicate weights – a form of the BRR technique.

## Example

```
/* Create Analytic Dataset for Repeated Cross-Section or Pooled Analysis */
/* Merge 2018 administrative records (HISUMRY) file with 2018 cross-sectional weights (CENWGTS) file */
data mcbs18;
    merge survey18.CENWGTS (in = a drop = VERSION)
          survey18.HISUMRY (keep = BASEID H_OPMDCD);
    by BASEID;
    if a;
run;

/* Merge 2017 administrative records (HISUMRY) file with 2017 cross-sectional weights (CENWGTS) file */
data mcbs17;
merge survey17.CENWGTS (in = a drop = VERSION)
      survey17.HISUMRY (keep = BASEID H_OPMDCD);
    by BASEID;
    if a;
run;

/* Append 2017 and 2018 cross-sectional files */
data mcbs_analytic_file;
    set mcbs17 mcbs18;
run;
```

## SAS

\* Estimate Percent of Dual 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 Dual 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

<sup>28</sup> 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.

```
svyset _n [pweight = ceyrswgt], brrweight(ceyrs001-ceyrs100) fay(.3) vce(brr)
```

\* Estimate Percent of Dual 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 Dual Eligible Medicare Beneficiaries (nationally representative, year-specific estimates)

```
svy brr, fay(.3): tab h_opmdcd surveyyr, 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 Dual 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 Dual Eligible Medicare Beneficiaries by Year (nationally representative, year-specific estimates)

```
prop.table(svytable(~H_OPMDCD + SURVEYYR, 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 * FLAG / 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 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 through subgroup analysis:

```
data mcbsdata_subset;
  set mcbsdata;
  if FLAG = 1 then output;
run;

proc surveyfreq data=mcbsdata_subset VARMETHOD = brr (fay=.30);
  table SEX * DIABETES / col notot;
  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 * FLAG / col notot;
  strata SUDSTRAT;
  cluster SUDUNIT;
  weight CEYRSWGT;

run;
```

## Appendix D: Initial Interview Variables

### Exhibit D.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 OTHRLANG WHATLANG
DEMO	Education	SPDEGRCV
DEMO	Income	INCOME

CHRNCOND	Reason for Medicare Eligibility	EMHBP EMMYOCAR EMCHD EMCFAIL EMHRTCND EMSTROKE EMCSKIN EMCANCER EMARTERY EMARTHRRH 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 E: Table of Links to MCBS Documentation

MCBS Resources	Links
CMS MCBS website	<a href="https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS">https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS</a>
MCBS LDS file information	<a href="https://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/Data-Disclosures-Data-Agreements/DUA_-_NewLDS">https://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/Data-Disclosures-Data-Agreements/DUA_-_NewLDS</a>
MCBS PUF	<a href="https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/MCBS-Public-Use-File/index">https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/MCBS-Public-Use-File/index</a>
CMS Chronic Conditions Warehouse (CCW)	<a href="https://www.ccwdata.org/web/guest/home/">https://www.ccwdata.org/web/guest/home/</a>
Data User's Guides, Methodology Reports and Codebooks	<a href="https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS/Codebooks">https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS/Codebooks</a>
Chartbook	<a href="https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS/Data-Tables">https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS/Data-Tables</a>
Early Look and Data Briefs	<a href="https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS/Data-Briefs">https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS/Data-Briefs</a>
Bibliography	<a href="https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS/Bibliography">https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS/Bibliography</a>
Questionnaire Specifications	<a href="https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS/Questionnaires">https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS/Questionnaires</a>