MCBS
Medicare Current Beneficiary Survey

2015 | METHODOLOGY REPORT

Centers for Medicare & Medicaid Services (CMS)
Office of Enterprise Data and Analytics (OEDA)
Date: February 16, 2018

TABLE OF CONTENTS

1. Introduction to MCBS 1

1.1 Purpose and Goals 1

1.2 Survey Overview and History 1

1.3 Key Data Products and Analyses 3

2. Changes Unique to 2015 5

2.1 Sampling 5

2.2 Data Collection 5

2.3 Questionnaires 6

2.4 Data Processing 7

3. Sample Design for the MCBS 2015 Panel 9

3.1 Overview of MCBS Sample Design 9

3.2 Selection of MCBS PSUs 12

3.3 Selection of MCBS SSUs 13

3.4 Selection of Beneficiaries for the 2015 MCBS Panel 14

Sample Size Determination 16

2015 Sampling Frame 22

Sample Selection for the 2015 Panel 24

Sampling Results 27

Coverage Analysis of the 2015 Sampling Frame 29

3.5 Continuing Sample (2012-2014 Panels) 31

3.6 Fielded Sample Sizes by Panel and Round 32

4. Instrument and Materials Design 33

4.1 Community Questionnaire Content 34

Interview Type 34

Community Questionnaire Flow 35

Core Section Content 38

Topical Section Content 46

Changes to the Community Questionnaire for 2015 47

4.2 Facility Instrument Content 50

Interview Type 50

Facility Screener 51

Facility Instrument Flow 51

Core Section Content 53

Changes to the Facility Questionnaire for 2015 57

4.3 CAPI and Case Management System Programming and Testing 57

Community Instrument 57

Facility Screener and Instrument 58

Case Management System 58

4.4 Letters and Other Respondent Materials 59

5. Interviewer Recruitment and Training 60

5.1 Interviewer Recruitment and Staffing 60

5.2 Interviewer Training Programs for 2015 60

Interviewer Training Goals and Objectives 60

In-Person Interviewer Training 60

Remote Interviewer Training 61

6. Data Collection 62

6.1. Clearance 62

OMB Approval 62

IRB Approval 62

6.2. Data Collection Process and Procedures 63

Data Collection Schedule and Timeline 63

Sample Releases and Preloads 63

Contacting Efforts and Outreach Rules 64

Beneficiary Eligibility for MCBS Survey 65

MCBS Data Collection Protocols 66

Crossover Definitions and Procedures 69

Proxy Interviews and Assistants 70

Interviewing Languages 71

Questionnaire Breakoffs 71

6.3 Data Collection Results 71

6.4 Data Collection and Quality Control 72

7. Data Processing and Data Delivery 73
7.1 Data Editing Overview... 73
Process Description 73
7.2 Preload Editing and File Production 75
Community and Facility Preload Timeline 76
Community and Facility Preload Process Description 76
7.3 MCBS 2015 Survey Data File.. 77
File Contents 77
7.4 MCBS 2015 Cost Supplement File.. 78
Eligible Population.. 78
Reference Period... 79
Insurance Timeline.. 79
Event Cost Consolidation 79
Prescription Medicine File.. 80
Facility Stay File 80

8. Weighting and Imputation 81
8.1 Overview 81
8.2 MCBS Weighting Procedures.. 81
Overview... 81
Process 81
2015 Initial Weighting 81
2015 Survey File Weights 87
2015 Cost Supplement Weights.. 92
2015 Topical Module Weights 93
8.3 MCBS Imputation Processes 96
Overview... 96
Income and Asset Imputation 96
Prescription Medicine and Non Prescription Medicine Imputation 97

9. Response Rates and Nonresponse 102
9.1 Response Rates.. 102
2015 Cost Supplement File Response Rates 102
2015 Survey File Response Rates 104

9.2 Nonresponse Bias Analysis 108

MCBS Attrition Rates *108*

Comparison of Respondents to Non-Respondents: Frame Characteristics *111*

Comparison of Respondents to Non-Respondents: Medicare Claims Payment Measures *117*

Comparison of Respondents to Non-Respondents: Chronic Condition Attributes *118*

Summary and Implications *119*

10. Using MCBS Data Files 120

10.1 MCBS Data User's Guide 120

10.2 MCBS Public Use Data File 120

10.3 MCBS Limited Data Sets 120

11. Glossary 121

TABLE OF EXHIBITS

Exhibit 1.2.1: MCBS Participation Timeline 2

Exhibit 1.3.1: 2015 Contents of Data Releases 4

Exhibit 2.3.1: 2015 and Typical MCBS Data Collection Year 6

Exhibit 2.3.2: 2015 Changes to Questionnaire Administration 7

Exhibit 3.1.1: 2010-2015 MCBS Rotating Panel Design 10

Exhibit 3.1.2: 2015 MCBS Sampling Strata 12

Exhibit 3.4.1: Assumed Rates (in Percent) Used in Determining Sample Sizes for the MCBS 2015 Panel, by Age Group 18

Exhibit 3.4.2: 2015 Fall Planning Sample Sizes and Resulting Projected Survey File and Cost Supplement File Estimates, by Age Group, Including Recent Enrollees and Excluding Hispanic Oversample 20

Exhibit 3.4.3: Expected Completed Interviews with Self-Reported Hispanic Beneficiaries, with and without Oversampling, by 2018 21

Exhibit 3.4.4: Number of Beneficiaries in 2015 Enrollment Data Subsample Extracts (Combined) and 2015 Sampling Frame, by Stratum 23

Exhibit 3.4.5: 2015 MCBS Panel, Final Sampling Fractions by Stratum 26

Exhibit 3.4.6: 2015 MCBS Panel, Number of Beneficiaries Selected by Age Group 27

Exhibit 3.4.7: 2015 MCBS Panel, Number of Beneficiaries Selected by Stratum 27

Exhibit 3.4.8: 2015 MCBS Panel, Number of Current-Year Enrollees Selected by Age Group .. 28

Exhibit 3.4.9: Number of Beneficiaries in 2015 Enrollment Data Subsample and Estimated 2015 MCBS Population, by Stratum 30

Exhibit 3.6.1: 2015 Fielded Sample Sizes by Round for Each Panel 32

Exhibit 4.1: MCBS Questionnaire Overview 33

Exhibit 4.1.1: Community Questionnaire Interview Types 35

Exhibit 4.1.2: 2015 MCBS Community Questionnaire Flow for Baseline Interview 36

Exhibit 4.1.3: 2015 MCBS Community Questionnaire Flow for Continuing Interview 37

Exhibit 4.1.4: 2015 MCBS Community Core Sections by Data File and Administration Schedule 39

Exhibit 4.1.5: Utilization and Cost Section Flow 44

Exhibit 4.1.6: 2015 MCBS Community Topical Sections by Data File and Administration Schedule 46

Exhibit 4.1.7: 2015 MCBS Topical Community Questionnaire Section Changes 48

Exhibit 4.2.1: Facility Instrument Interview Types 50

Exhibit 4.2.2: 2015 MCBS Facility Instrument Flow for Baseline Interview 52

Exhibit 4.2.3: 2015 MCBS Facility Instrument Flow for Continuing and Crossover Interviews .53

Exhibit 4.2.4: Facility Core Sections by Data File and Administration Schedule 54

Exhibit 6.2.1: Typical MCBS Data Collection Round Case Release Schedule 63

Exhibit 6.2.2: MCBS Contacting Guidelines by Interview Round 64

Exhibit 6.2.3: Interviewer-Reported Abstraction Rates by Facility Instrument Section....................69

Exhibit 6.3.1: 2015 Completed Interviews by Component....................72

Exhibit 7.1.1: Data Review and Editing Process 74

Exhibit 8.2.1: 2015 MCBS Data Files Summary of Weights 83

Exhibit 8.2.2: Control Totals for Ever-Enrolled Weight Raking, Dimension 1: Age Group........91

Exhibit 8.2.3: Control Totals for Ever-Enrolled Weight Raking, Dimension 2: Enrollment Year 92

Exhibit 8.2.4. 2015 Data Year Topical Module Survey Weights Datasets and Contents 94

Exhibit 8.3.1. Payers and Payment Amounts Missing Together, Total Payment Known 98

Exhibit 8.3.2. Payers and Payment Amounts Missing Differentially, Total Payment Known 99

Exhibit 8.3.3. Total Payment Unknown....................99

Exhibit 9.1.1: 2015 MCBS Annual Cost Supplement File Unconditional Response Rates 104

Exhibit 9.1.2: 2015 MCBS Annual Cost Supplement File Conditional Response Rates...........104

Exhibit 9.1.3: 2015 MCBS Annual Survey File Unconditional Response Rates for Ever-Enrolled Beneficiaries 106

Exhibit 9.1.4: 2015 MCBS Annual Survey File Unconditional Response Rates for Continuously-Enrolled Beneficiaries 107

Exhibit 9.1.5: 2015 MCBS Annual Survey File Conditional Response Rates for Ever Enrolled Beneficiaries 107

Exhibit 9.1.6: 2015 MCBS Annual Survey File Conditional Response Rates for Continuously-Enrolled Beneficiaries 108

Exhibit 9.2.1: Unweighted Attrition Rates by Panel, Fall 2015 to Winter 2017: 2015 Survey File 109

Exhibit 9.2.2: Unweighted Attrition Rates by Panel, Fall 2015 to Winter 2017: 2015 Survey File 111

Exhibit 9.2.3: 2015 Panel Respondents vs. Non-respondents in Fall 2015, by Frame Characteristics 112

Exhibit 9.2.4: 2012-2014 Panel Respondents vs. Non-respondents in Fall 2015, by Frame Characteristics 114

Exhibit 9.2.5: 2015 Panel Non-respondents (NR) vs. Hard-to-Contact Respondents (HR) and Other Respondents (OR), by Frame Characteristics: Fall 2015 115

Exhibit 9.2.6: Logistic Regression Model of 2015 Panel Non-respondents, Fall 2015 117

Exhibit 9.2.7: 2015 Claims Payment Measures for 2015 Panel[a] Non-Respondents (NR), Hard-to-Contact Respondents (HR), and Other Respondents (OR): Fall 2015 118

Exhibit 9.2.8: Fall 2015 Respondents vs. Non-Respondents, by Round, Across Chronic Condition Attributes 119

1. INTRODUCTION TO MCBS

1.1 Purpose and Goals

Medicare is the nation's health insurance program for persons 65 years and older and for persons younger than 65 years who have a qualifying disability. The Medicare Current Beneficiary Survey (MCBS) is 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 the Medicare program. A leading source of information on Medicare and its impact on beneficiaries, the MCBS provides important information on beneficiaries that is not 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 covering the population of beneficiaries, including the population of Medicare beneficiaries over age 65 as well as beneficiaries with disabilities under age 65, in the United States and its territories. Fieldwork for the first round of data collection began in September 1991; since then, it has continued to collect and provide essential data on the costs, use, and health care status of Medicare beneficiaries. Recently celebrating its 25th anniversary of continuous data collection, the MCBS has completed 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, including the creation of new benefits such as Medicare's Part D prescription drug benefit.

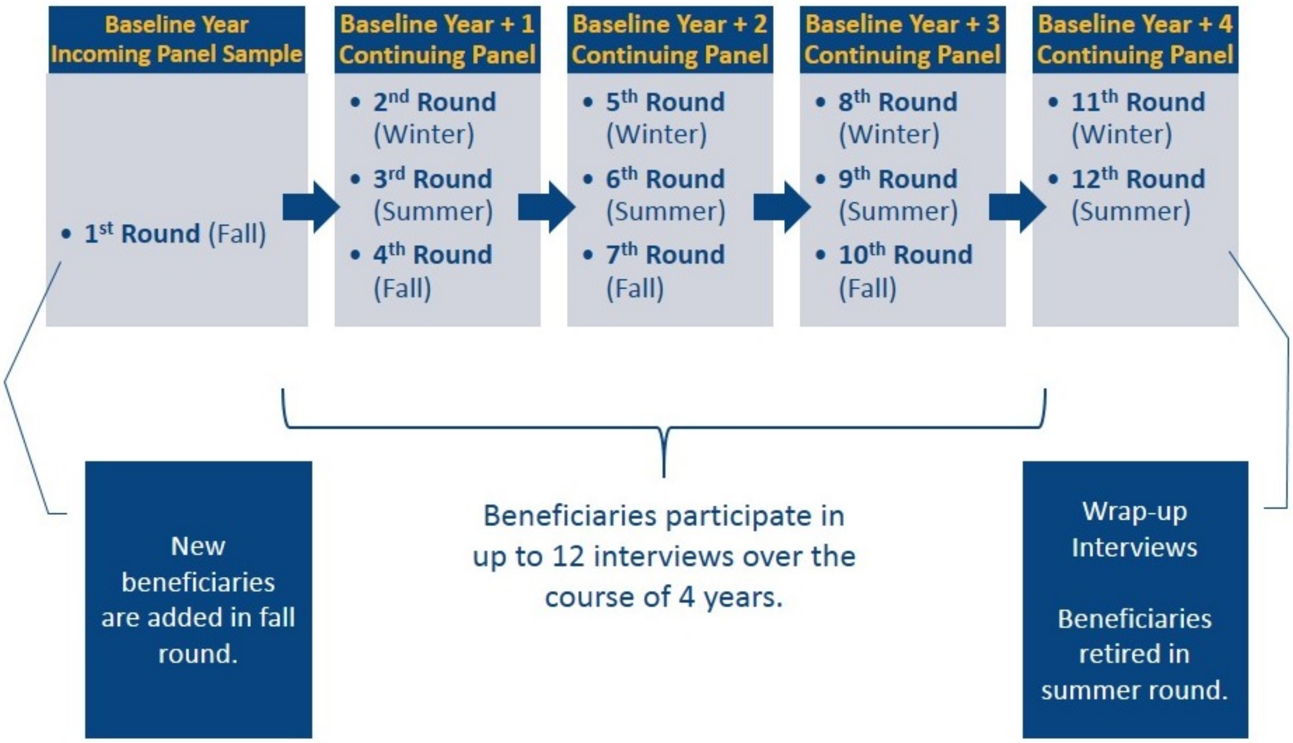
1.2 Survey Overview and History

In its initial design, the MCBS was to serve as a traditional longitudinal survey of the Medicare population, with no predetermined limit to the duration of participation. However, beginning in 1994, participation of beneficiaries in the MCBS was limited to no more than four years.

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 the re-interviewing of the same respondent (or appropriate proxy respondents) until they have completed four years of participation (up to 12 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 1.2.1 depicts the timeline of participation for beneficiaries selected to be in the MCBS sample.

Exhibit 1.2.1: MCBS Participation Timeline[1]



Detailed information on the sampling design can be found in Chapter 3 of this report. Chapter 6 describes the data collection fielding procedures, including eligibility for each round of the interview, and Chapter 9 summarizes the results of data collection, including response rates and an analysis of nonresponse bias.

---

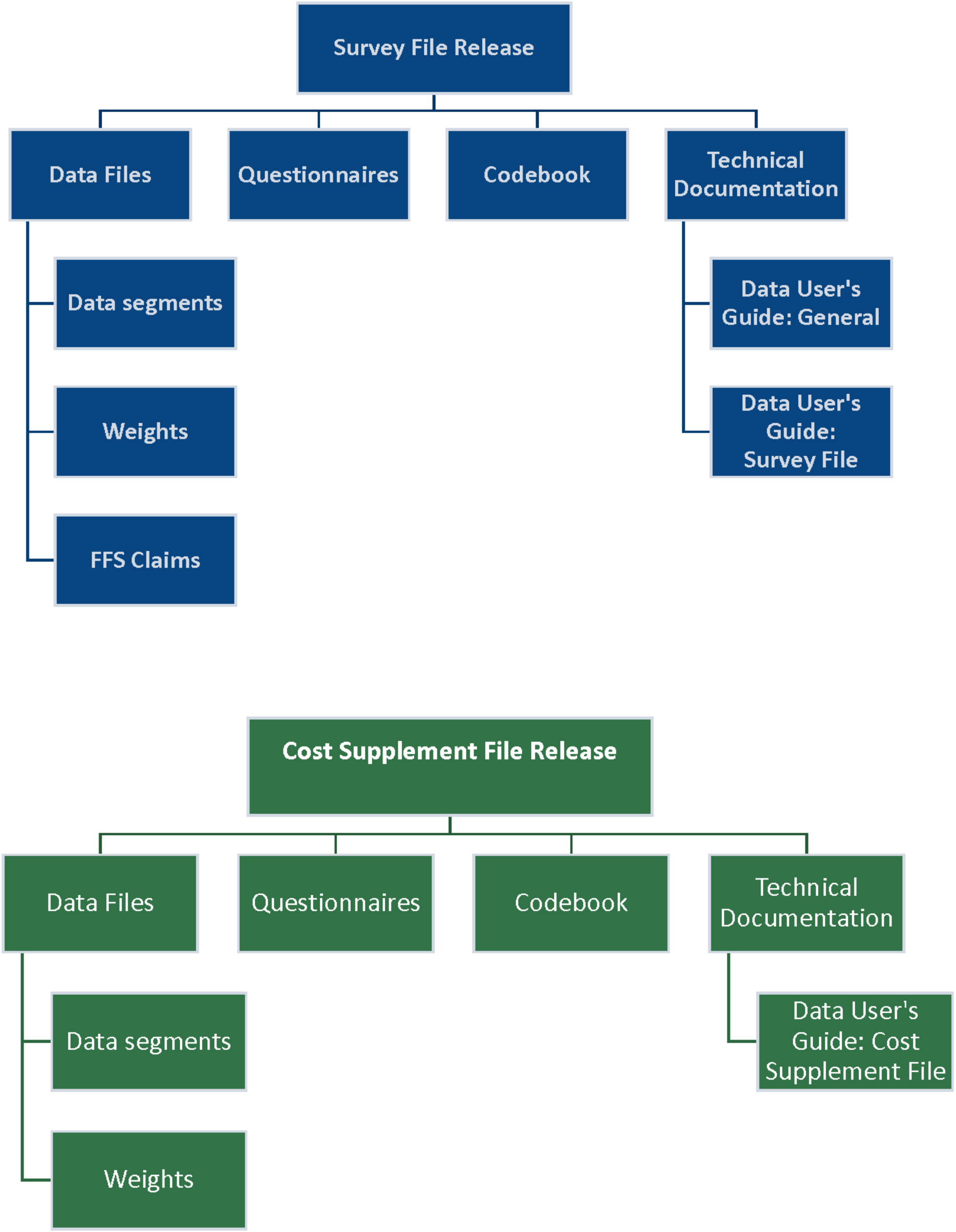
[1] Note that the 2015 data year had a modified schedule. See Exhibit 2.3.1.

1.3 Key Data Products and Analyses

MCBS data are made available via releases of annual files. For 2015, two annual Limited Data Set (LDS) releases, the Survey File and the Cost Supplement File, and one PUF (based on the Survey File data only) are released. The LDS releases each contain multiple files, called segments, which are easily linkable through a common beneficiary key ID. 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 the data release-specific chapters of the MCBS Data User's Guide (see MCBS Data User's Guide – Survey File and MCBS Data User's Guide – Cost Supplement File).

Exhibit 1.3.1 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 Fee-for-Service claims data, which provide CMS administrative information on medical services and payments paid by Medicare claims; 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 in order to link cost and utilization variables with demographic or health insurance coverage variables.

Exhibit 1.3.1: 2015 Contents of Data Releases



Chapter 4 of this report provides information on the specific questionnaire sections associated with each data file. Chapter 7 describes the creation of these data files and Chapter 8 provides an overview of weighting and imputation procedures. 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 can be found in the data file-specific chapters of the MCBS Data User's Guide (see MCBS Data User's Guide – Survey File and MCBS Data User's Guide – Cost Supplement File).

2. CHANGES UNIQUE TO 2015

Several key changes were made to the MCBS during 2015, affecting the areas of sampling, data collection, questionnaire design, and data processing, including weighting and imputation procedures. These changes are highlighted below and described later in this report.

2.1 Sampling

Sample eligibility: 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 MCBS panel. Prior to 2015, only beneficiaries who became eligible on or before January 1 of the sampling year were eligible to be sampled. Including the current-year eligible beneficiaries allows for the release of data products up to one year earlier. Additional information can be found in Section 3.4.

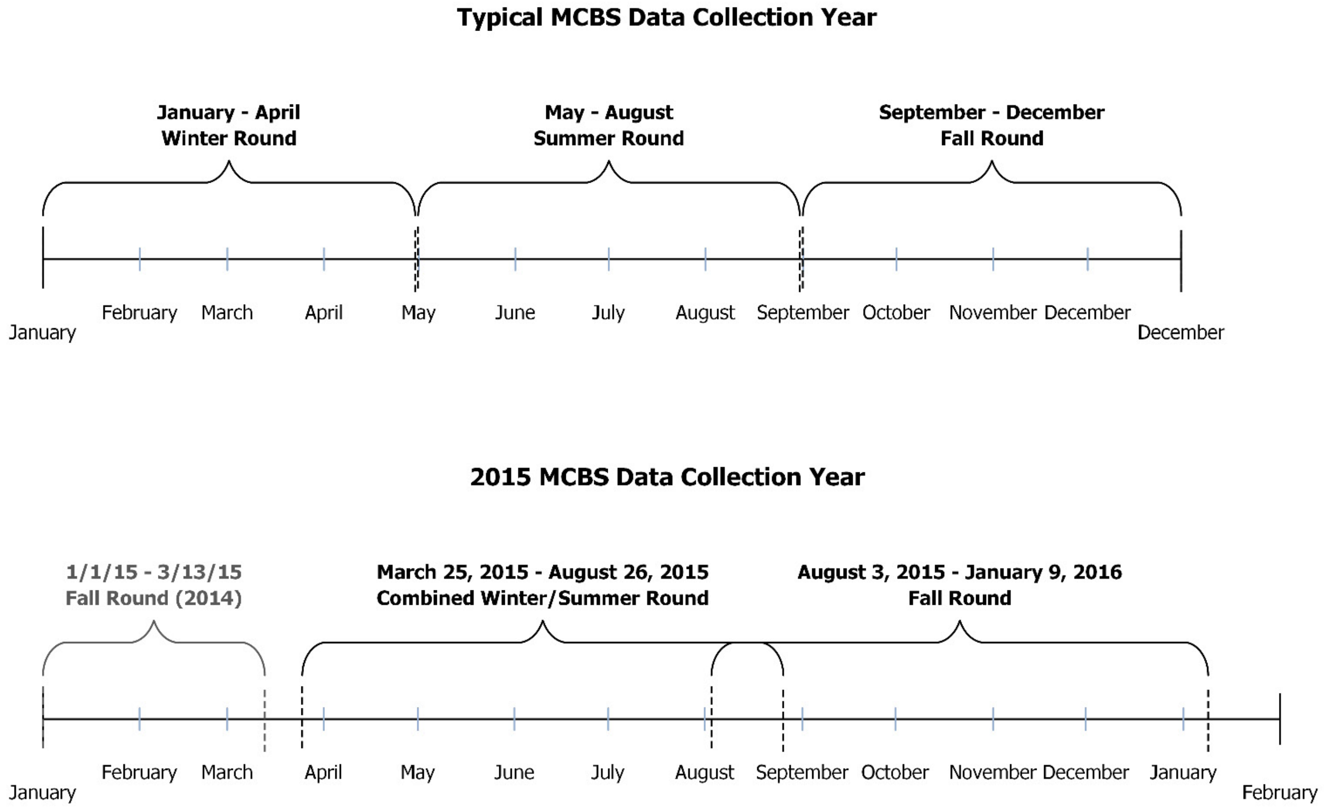
Census tracts replaced ZIP Code areas for Secondary Sampling Units (SSUs): Beginning in Fall 2014, Census tracts or groups of tracts, replaced ZIP Code areas as SSUs for the new panel selected and fielded each fall. See Section 3.3 for more information.

Hispanic oversample: Beginning in 2015, Hispanic beneficiaries living outside of Puerto Rico were oversampled for the MCBS. The main goals of the oversampling were to increase the number of Hispanic beneficiaries to allow for more precise estimates of health disparities experienced by these populations and to increase the proportion of MCBS Hispanic beneficiaries from outside Puerto Rico.

2.2 Data Collection

In 2015, the traditional model of conducting the MCBS in three rounds was modified. Due to a transition period between contractors that began in 2014 and continued into 2015, additional time was required to complete the re-programming of all questionnaire sections. Therefore, the Fall 2014 round was conducted from September 2014 through March 2015; the Winter 2015 and Summer 2015 rounds were combined into a single, longer data collection period that was conducted from March 2015 through August 2015; and the Fall 2015 round was then conducted from August 2015 through early January 2016. Thus, the 2015 MCBS data releases reflect data collected from March 2015 through the first week in January 2016 (see Exhibit 2.3.1).

Exhibit 2.3.1: 2015 and Typical MCBS Data Collection Year


Typical MCBS Data Collection Year

2015 MCBS Data Collection Year

2.3 Questionnaires

Due to the 2015 reprogramming of the instrument in data collection in 2015, the rotation of some Community questionnaire sections was also affected. Specifically, the Beneficiary Knowledge and Information Needs Supplement (KNQ), usually asked in the winter, and the Drug Coverage (RXQ), usually asked in the summer, were included in Winter/Summer 2015; Income and Assets Supplement (IAQ), usually asked in the summer, was shifted to Fall 2015 (Round 73); and Patient Activation (PAQ), usually asked in the summer, was dropped in 2015 due to similar content in the revised Usual Source of Care that included new Patient Perceptions of Integrated Care (USQ-PPIC) measures asked in Fall 2015. Exhibit 2.3.2 summarizes these changes. Additional details of questionnaire content and changes made in 2015 are found in Section 4.1.

Exhibit 2.3.2: 2015 Changes to Questionnaire Administration

Questionnaire Section	Typical Season of Administration	Season of Administration in 2015
Beneficiary Knowledge and Information Needs (KNQ)	Winter	Winter/Summer
Drug Coverage (RXQ)	Winter	Winter/Summer
Income and Assets (IAQ)	Summer	Fall
Patient Activation (PAQ)	Summer	Dropped

2.4 Data Processing

Naming conventions and reorganization: Starting with the 2015 MCBS, the two annual releases are now referred to as follows:

- Survey File (formerly Access to Care): as before, this file contains survey data augmented with administrative and claims data to allow for analysis regarding beneficiaries' self-reported health status, health conditions, access to health care and satisfaction with health care.
- Cost Supplement File (formerly Cost and Use): as before, provides cost and utilization data and can be linked to the Survey File to conduct analysis on healthcare cost and utilization. Of note, previously demographic and survey data were included in this file; but now only the MCBS Survey File exclusively contains these data. Users will require both files to conduct analyses.

Additionally, data file segments within the release are no longer referred to as RICs,[2] but are now titled according to the topic of data included in the file. More complete descriptions of the data file contents can be accessed in the MCBS Data User's Guide.

Data file structure and weights: For analysis of Survey File data, there are two available populations of inference that can be obtained through the use of two distinct weights. First, the ever enrolled Survey File weights are populated for the largest group of respondents and represent the population of beneficiaries who were entitled and enrolled in Medicare for at least one day at any time during the 2015 calendar year. Second, the always enrolled Survey File weights are populated for, and represent the population of, beneficiaries who were enrolled from the first of the year (January 1, 2015) through completion of their fall interview. These weights are identical in terms of estimation to the historical Access to Care (ATC) weights that were available in previous years, and can be used in the same way as the prior weights.

---

[2] RIC stands for Record Identification Code

Analyses of the Cost Supplement File data are conducted with the Cost Supplement weights, which represent an ever enrolled population during 2015 identical to the population represented by the ever enrolled Survey File weights but are populated for a smaller subset of respondents with complete cost and utilization data. Joint analysis of both Survey File and Cost Supplement File data should be conducted using these Cost Supplement File weights. Detailed information about the data file structure and weights can be found in Section 8.3.

Data editing and cleaning procedures: MCBS data files receive thorough editing and cleaning prior to release. Quality control checks are conducted to confirm each analytic file is structurally sound. These checks include confirming that all necessary variables are present, checking variable attributes, and identifying high rates of missing data.

Logic and reasonableness checks were also implemented for each analytic file. Logic checks verify that the questionnaire worked as expected, particularly with respect to questionnaire routing, and reasonableness checks identify values that are not explicitly disallowed by the questionnaire. Based on the results of this data review, new edits were developed to correct the errors during data cleaning. Additional information about data editing can be found in Chapter 7.

Imputation methods changes: Due to the substantial changes in the Income and Assets Questionnaire (IAQ), the 2015 IAQ imputation imputes a new set of IAQ variables. New imputation methods were tested and used for this new set of variables. For the 2015 non-prescription medicine imputation, methodological changes were made that improve the efficiency and accuracy of the imputation. See Section 8.3 for more detail about imputation procedures.

3. SAMPLE DESIGN FOR THE MCBS 2015 PANEL

3.1 Overview of MCBS Sample Design

The MCBS employs a three-stage cluster sample design. In 2015, the survey continued to use the set of 107 primary sampling units (PSUs) that have been employed for sampling for the MCBS since 2001 as well as the census tract-based secondary sampling units (SSUs) that were selected in 2014[3]. At the third stage, Medicare beneficiaries, the ultimate sampling units (USUs), were selected from within the selected tract-based SSUs. The MCBS sample design underwent a few significant changes in 2015, including the oversampling of Hispanic beneficiaries and the inclusion of current-year enrollees in the sampling frame.

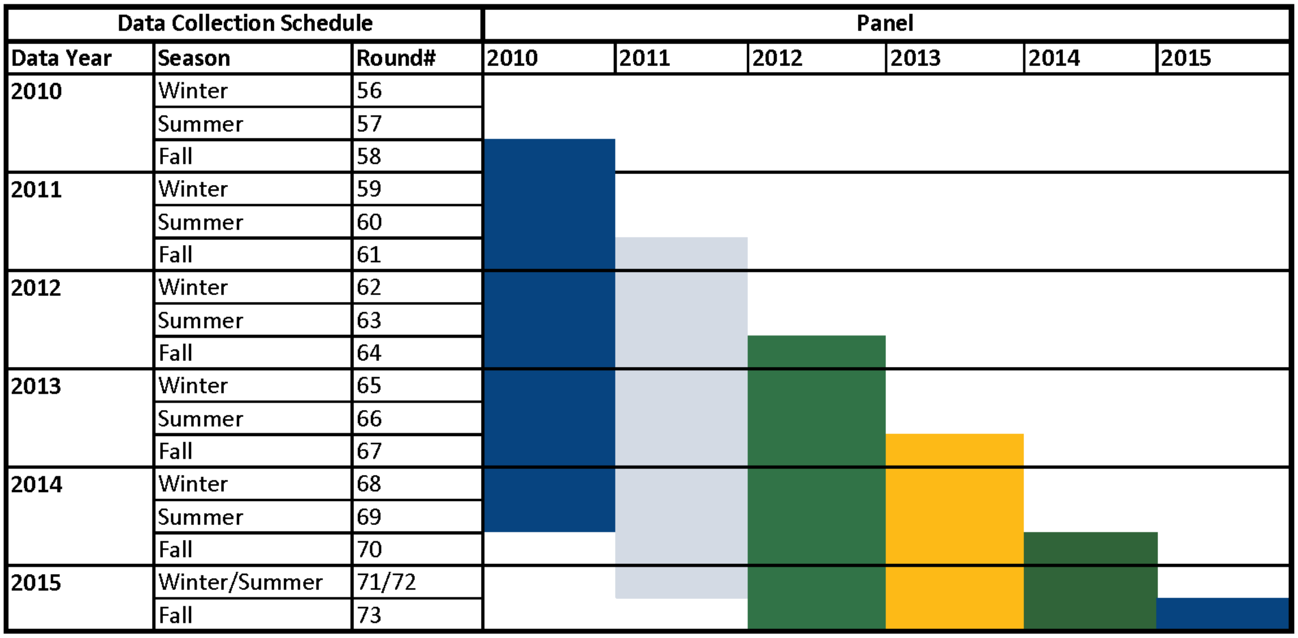
In 2015, MCBS continued to use the sample rotation pattern used historically. In particular, the newly selected 2015 Panel, and the panels selected in 2012, 2013, and 2014, continued into Fall 2015 and beyond according to their established rotation schedules.[4] The 2011 Panel (which was first fielded in Fall 2011) exited in Winter/Summer 2015 and was replaced with a new sample of beneficiaries in Fall 2015. Exhibit 3.1.1 displays the MCBS rotating panel design from 2010 to the present panel.

---

[3] These SSUs were sized to last twenty years; therefore, there is no current plan to reselect the SSUs.

[4] A new panel is added each Fall and retains the year of its entry as its sampling panel designation for projections and response rate analysis. Once a panel is selected, it remains in the MCBS for four years, participating in twelve total rounds.

Exhibit 3.1.1: 2010-2015 MCBS Rotating Panel Design



Data Collection Schedule			Panel					
Data Year	Season	Round#	2010	2011	2012	2013	2014	2015
2010	Winter	56						
	Summer	57						
	Fall	58	■					
2011	Winter	59	■					
	Summer	60	■					
	Fall	61	■	■				
2012	Winter	62	■	■				
	Summer	63	■	■				
	Fall	64	■	■	■			
2013	Winter	65	■	■	■			
	Summer	66	■	■	■			
	Fall	67	■	■	■	■		
2014	Winter	68	■	■	■	■		
	Summer	69	■	■	■	■		
	Fall	70		■	■	■	■	
2015	Winter/Summer	71/72		■	■	■	■	
	Fall	73			■	■	■	■

This section documents the procedures used to select the new sample for Fall 2015 (i.e., the 2015 Panel). The 2015 Panel will be retained in the study for the four years specified under the MCBS sample rotation scheme and is designed to: (a) replace approximately one-third[5] of the respondents in the existing MCBS sample; and (b) extend survey coverage to persons added to the Medicare roles during the previous year (and, beginning in 2015, during the current year, as well; see Section 3.4 for details).

PSUs. To maintain continuity with the past, the set of PSUs for the 2015 Panel consisted of the 107 previously constructed PSUs used for MCBS sampling since 2001, including 29 certainty PSUs and 78 non-certainty PSUs.

SSUs. Prior to 2014, MCBS SSUs were made up of ZIP Codes and ZIP Code fragments. Beginning in 2014, the survey implemented the use of census tracts, or clusters of adjacent tracts, as SSUs. A total sample of 703 SSUs was selected within the 107 PSUs, consisting of a proportional allocation of 242 SSUs to the 29 certainty PSUs and an equal allocation of approximately 6 SSUs to each of the 78 non-certainty PSUs. These SSUs are sized to last for a full twenty years of use for MCBS sampling. 2015 is the second year in which the annual MCBS panel sample is selected

---

[5] Due to the cumulative effects of attrition over time as well as cost-related sample cuts from past years, the number of MCBS respondents varies by panel, with fewer respondents in the older panels than in newer ones. Thus, while the newly-selected panel replaces one of four existing panels, the net effect has been to replace about one-third of the existing MCBS respondents. Furthermore, because attrition has been higher than expected in recent years, some of the newer panels may be required to replace more than one-third of the respondents.

from within the new SSUs; it will take four years (until the 2017 Panel is selected) for the MCBS sample to fully transition from the ZIP-based to the census tract-based SSUs.

USUs. The third stage of sampling was the selection of Medicare beneficiaries from within each SSU. Previously, to be eligible for sample selection, beneficiaries had to be eligible for Medicare and enrolled by the first day of the sampling year, January 1, 2015. Beginning with the 2015 Panel, however, all current-year enrollees are also eligible to be sampled as part of a new sampling approach. As a result, all beneficiaries who are enrolled in Medicare before January 1, 2016 are eligible to be sampled as part of the 2015 Panel. Sampling of current-year enrollees is discussed in detail in Section 3.4.

In addition to the inclusion of current-year enrollees in the sampling frame for the 2015 Panel, Hispanic beneficiaries living outside of Puerto Rico were oversampled beginning in 2015. An additional 75 completed interviews from Hispanic beneficiaries are targeted per year with the aim of producing 1,500 Hispanic beneficiaries in the 2018 Survey File. Oversampling of Hispanic beneficiaries is discussed in detail in Section 3.4.

The sampling frame for the Medicare beneficiaries begins with Medicare administrative enrollment data. To avoid duplication in the various panels of MCBS beneficiaries, a unique and disjoint 5-percent sample of the enrollment data is specified annually by CMS for the MCBS. The most recent 5-percent file was used as a basis for selecting the sample for the 2015 MCBS Panel. A first extract of the enrollment data 5-percent file was provided in April 2015, and the bulk of the 2015 Panel sample was selected from that extract. Two additional extracts of the enrollment data 5-percent file, containing only new enrollees who were not included in the initial extract, were also needed to support sampling of current-year enrollees.[6] The combination of these extracts constitutes the full frame from which the 2015 Panel was selected. Details about the sampling frame construction can be found in Section 3.4.

The MCBS enrollment data 5-percent file extracts were subset based on eligibility and other criteria (described in detail later in this section) and then geocoded to the tract level. The set of all records that geocoded to the selected SSUs constituted the MCBS sampling frame of beneficiaries. A random sample of beneficiaries residing in the selected SSUs was then selected within defined age group and ethnicity (Hispanic/non-Hispanic) strata within the U.S., and within age group strata in Puerto Rico. An ethnicity flag (see Section 3.4 for a full description) was used to classify beneficiaries into the Hispanic strata; a value of "yes" indicates that the beneficiary is expected to be Hispanic; a value of "no" indicates that the beneficiary is not expected to be Hispanic. (Actual, or self-reported, Hispanic origin status may differ from the ethnicity flag.) Thus, the sample was selected within the strata displayed in Exhibit 3.1.2.

---

[6] Note that while all new enrollees added to the enrollment data since the previous extract(s) are received, only new *current-year* enrollees are sampled.

Exhibit 3.1.2: 2015 MCBS Sampling Strata

U.S. Hispanic	U.S. Non-Hispanic	Puerto Rican Resident
Under 45 U.S. Hispanic	Under 45 U.S. non-Hispanic	Under 45 Puerto Rican resident
45 - 64 U.S. Hispanic	45 - 64 U.S. non-Hispanic	45 - 64 Puerto Rican resident
65 - 69 U.S. Hispanic	65 - 69 U.S. non-Hispanic	65 - 69 Puerto Rican resident
70 - 74 U.S. Hispanic	70 - 74 U.S. non-Hispanic	70 - 74 Puerto Rican resident
75 - 79 U.S. Hispanic	75 - 79 U.S. non-Hispanic	75 - 79 Puerto Rican resident
80 - 84 U.S. Hispanic	80 - 84 U.S. non-Hispanic	80 - 84 Puerto Rican resident
85 and over U.S. Hispanic	85 and over U.S. non-Hispanic	85 and over Puerto Rican resident

Sampling rates varied by stratum, with the younger beneficiares with disabilities (under 45) and elderly beneficiaries (85 and older) strata being oversampled. Hispanics are also oversampled relative to their non-Hispanic age stratum counterparts.[7] The MCBS sampling design for an annual panel provides nearly self-weighting (i.e., equal probabilities of selection) samples of beneficiaries within each of the 21 sampling strata.

The sample size for the 2015 Panel was determined to achieve approximately 11,500 responding beneficiaries across all panels that would comprise the 2018[8] Cost Supplement File. It should be noted that the Hispanic oversample is based on the Survey File counts because it is an annual addition to the fall interview.[9] The targeted annual completed sample size of 11,500 responding beneficiaries with Cost Supplement data is expected to be comprised of approximately 1,000 beneficiaries from each of the under 65 (disability) age groups and approximately 1,800-2,000 beneficiaries from each of the remaining age groups.

3.2 Selection of MCBS PSUs

The original MCBS PSU sample was selected in 1991 using a sampling frame that was developed using 1980 Census data. In 2001, the set of PSUs was redesigned and reselected. For the redesign, those PSUs in the continental U.S. with at least 224,000 Medicare beneficiaries, according to the June 2000 enrollment data 5-percent file, were included in the sample with certainty. The cutoff of 224,000 corresponds roughly to a probability of selection of 0.75 under a PPS (probability-proportionate-to-size) sample design. The use of the specified cutoff resulted in designating the 28 largest PSUs in the continental U.S. as certainty PSUs. Of these, 27 were also

---

[7] The Hispanic oversample is relatively small and does not impact the sampling rates as much as the age group targets do.

[8] The target corresponds to the 2018, rather than the 2015, Cost Supplement File because 2018 is the final year that the 2015 panel beneficiaries will contribute to a Cost Supplement File. The goal is to start with a large enough sample to achieve, after attrition and deaths, the required number of completes in the panel's final Cost Supplement year.

[9] While the base sample sizes are calculated with a Cost Supplement File target in mind, the oversamples were calculated based on a Survey File goal. This is because the oversamples are built over time via additional completes from the newest panel each year, which is easiest to enumerate in terms of Survey File yield.

certainties in the original design, and one had been selected as a noncertainty PSU in the original design. In addition, the largest PSU in Puerto Rico (the central part of the San Juan Metropolitan Statistical Area (MSA)) was included in the sample with certainty.

The remaining noncertainty PSUs were grouped by census region and MSA status (where Puerto Rico was treated as a separate "region" for sampling purposes). Within each major group of PSUs, detailed sampling strata were formed by sorting PSUs by the percentage of Medicare beneficiaries enrolled in Health Maintenance Organization (HMO) plans (and in some cases also by the percentage of Medicare beneficiaries who were minorities) and then forming strata of roughly equal size from this sorted list. Thirty-eight noncertainty strata were formed within the continental U.S., and one was formed in Puerto Rico. Two PSUs were then selected from each stratum with probabilities proportionate to size using procedures designed to maximize overlap with the existing MCBS sample. The procedure developed by Ernst (1986)[10] was used for this purpose. In the Ernst approach, each stratum in the new design is treated as a separate linear programming problem, and the optimization of the linear programming problem determines a set of coefficients that are used to select the new sample.

The PSUs are examined periodically for representativeness to the national Medicare population. The most recent analysis was conducted in 2016, and it was determined that a reselection of PSUs was not necessary at that time.

3.3 Selection of MCBS SSUs

Prior to 2014, MCBS SSUs were defined as clusters of ZIP Codes and ZIP Code fragments. In 2014, the MCBS SSUs were reselected using census tracts. This change reduces the need for maintenance of SSUs previously required due to the nature of ZIP Code boundary changes and allows for easier merging of MCBS data with U.S. Census Bureau data and other aggregate level geographic or environmental extant data.

The creation and selection of SSUs involved several steps. First, the enrollment data[11] were geocoded to the tract level. Next, enrollment data counts by tract were used to calculate SSU measures of size. Small tracts were combined to create SSUs that met a minimum measure of size. (The SSUs were sized to provide for a full twenty years of sampling for the MCBS.) Finally, a core sample of 703 SSUs, plus an additional reserve sample of 339 SSUs, were selected from a frame of 24,212 SSUs within the 107 PSUs using systematic probability proportional to size selection. A fixed number of SSUs was selected from each of the 78 noncertainty PSUs (subject to

---

[10] Ernst, L. (1986). Maximizing the overlap between surveys when information is incomplete, *European Journal of Operational Research*, 27, 192-20.

[11] Because the enrollment data are so plentiful and the geocoding process so laborious, only enrollment data records falling into ZIP Codes that overlapped with one of the 107 PSUs were geocoded to the tract level. Then, only those falling into one of the 107 PSUs were kept for the selection of SSUs.

a maximum of 6), and the remainder were proportionally allocated to the certainty PSUs (subject to a minimum of 6 and constrained to an even number).

A total of 703 core SSUs, comprised of 242 SSUs from the certainty PSUs and 461 SSUs from the noncertainty PSUs, was selected in 2014. An additional reserve sample of 339 SSUs (122 from certainty PSUs and 217 from the noncertainty PSUs) was also selected to provide the MCBS with the possibilities to expand the sample or to study special rare populations in future years. In 2015, the set of 703 core SSUs was used again for sample selection.

3.4 Selection of Beneficiaries for the 2015 MCBS Panel

The third stage of sampling is the selection of Medicare beneficiaries from within each SSU. In 2015, two major design innovations were introduced at this stage. First, current-year Medicare enrollees were now eligible to be sampled as part of each new annual panel. Their inclusion will allow for the release of data files up to one year earlier than previously possible.[12] Second, Hispanic beneficiaries were oversampled. These innovations are discussed in detail below.

Current-Year Enrollee Sample. Historically, to be eligible for sample selection, beneficiaries had to be eligible for Medicare and enrolled by January 1st of the sampling year (*t*), which resulted in the release of data products containing information about the cost and use of health care services (i.e., the Cost Supplement File) in reference year *t* during the middle of the year two years later (year *t* + 2).[13] Such late release arose because the year *t* cohort of beneficiaries, which contributes to the cost and use of health care services in reference year *t*, was not even sampled until year *t* + 1 and not initially interviewed until the fall round of year *t* + 1.

Beginning in 2015, the year *t* cohort[14] of beneficiaries (i.e., the set of current-year enrollees) was included in the sampling frame of beneficiaries from which the year *t* panel[15] was selected and resulted in the collection of health care cost and utilization (Cost Supplement File) data for these cases one year earlier than previously possible. This will allow for processing of these data to occur one year earlier than under the previous design, and may therefore allow for release of data files up to one year earlier.

The inclusion of the current-year enrollees introduced the need for multiple enrollment data extracts for sampling and multiple sample draws. Because not all 2015 enrollees are included in the enrollment data by the time the initial sampling needs to occur, additional extracts, or

---

[12] Persons who became eligible for Medicare during 2015 could have incurred health care costs in 2015. By including such persons in the sampling process up to a year earlier than was done previously, they can be appropriately represented in the 2015 Cost Supplement File up to a year earlier.

[13] Final versions of Medicare claims for a calendar year are available six months after the start of the subsequent calendar year.

[14] An annual cohort is the set of beneficiaries that are enrolled in Medicare and appear in the Medicare enrollment data within a given year.

[15] An annual panel is the set of beneficiaries sampled in a given year and initially interviewed in the fall round of that year.

"updates" to the original enrollment data extract for the 2015 Panel, were required. The first, and largest, extract, which contained the bulk of the 2015 sampling frame, was delivered by CMS in April 2015. The majority of the 2015 Panel was selected from this initial extract. Additional enrollment data extracts of 2015 enrollees were delivered in early August and late September 2015, and additional samples of 2015 enrollees were drawn from these extracts. The sampling frame for the 2015 Panel is made up of the beneficiaries in the three extracts falling into the MCBS PSUs and SSUs. A fourth and final extract was delivered in mid-January 2016 and used to fully enumerate the 2015 population of Medicare enrollees. Because data collection had already ended for Fall 2015, no sample was drawn from the January extract; however, the information was used for weights calibration. Please see the Coverage Analysis section below for a detailed description of this extract and the results of the coverage analysis.

Timing of the Interview. Members of the year *t* cohort of beneficiaries sampled under the new design will all be enrolled in Medicare sometime during sampling year *t*. Because these individuals may be more cooperative after they become eligible and have a connection to Medicare, and because the interview is geared toward those who are already enrolled, these sampled individuals are interviewed only after they are enrolled. The majority become eligible and enroll before Fall interviewing begins; for those not enrolled until after September 1, an interview is conducted with the sampled beneficiary after he or she enrolls in Medicare (i.e., on or after their enrollment date in the enrollment data).

Hispanic Oversample. Beginning in 2015, Hispanic beneficiaries living outside of Puerto Rico were oversampled in the new panel. Hispanics were identified and the sampling frame was stratified using a flag provided by CMS based on Census records of Hispanic surnames and other enrollment information such as language preference, and the Hispanic stratum was oversampled relative to the non-Hispanic stratum. The main goal of the oversampling was to increase the number of beneficiaries of Hispanic, Latino/a, or Spanish origin in the MCBS to allow for precise estimates of health disparities experienced by these populations by increasing the proportion of MCBS Hispanic beneficiaries from outside Puerto Rico. An additional 75 completed interviews with Hispanic beneficiaries are targeted annually with the goal of achieving 1,500 annual Hispanic completes across all panels by 2018.

Sample Selection Overview. The sample of MCBS beneficiaries was selected using systematic sampling within each PSU, and specifically only within the 703 core SSUs selected within the PSUs. In June, the majority of the 2015 Panel was selected along with a reserve sample of beneficiaries. (The reserve sample was included to give the opportunity of releasing additional sample in case initial planning assumptions proved to be incorrect. In the end, the reserve sample was not used.) In August and October, additional small samples[16] of 2015 enrollees were selected using the same sampling rates as for the initial sample.

For the 2015 MCBS Panel, an initial sample of 8,426 beneficiaries (including the Hispanic oversample) was selected in June. In August, an additional 144 current-year enrollees were

---

[16] Plus corresponding reserves, for use if necessary.

selected using the sampling rates computed for the first extract and were added to the 2015 MCBS Panel. In October, a further 65 current-year enrollees were selected, again using the sampling rates computed for the first extract, and added to the 2015 MCBS Panel. As of October, the 2015 MCBS Panel was complete, with a total of 8,635 beneficiaries (including the Hispanic oversample).

Details of the determination of the sample size, the construction of the sampling frame, and the selection of the sample of beneficiaries for the 2015 MCBS Panel are given below.

Sample Size Determination

The sample size requirements for the 2015 Panel were derived using estimated sample losses due to "immortals," deaths, and nonresponse. Immortals are defined to be (a) persons in the CMS sampling frame who enrolled prior to the year preceding the sampling year and are determined to be deceased at the first or second interview, and whose date of death is confirmed by a proxy to be prior to the sampling year but for whom no death is recorded in CMS administrative updates; (b) persons in the CMS sampling frame who enrolled prior to the year preceding the sampling year and are determined to be ineligible for Medicare in the first or second interview, and whose loss of entitlement is confirmed by the respondent or a proxy to be prior to the sampling year but for whom there is no record of having lost eligibility in CMS administrative updates; or (c) persons who enrolled prior to the year preceding the sampling year and died or lost Medicare eligibility prior to the sampling year based on CMS administrative updates. These three types of immortals all share the characteristic that they would never have been sampled if up-to-date and accurate information on death and eligibility status had been available in the CMS sampling frame.[17] Sampled beneficiaries who were deceased at the first or second interview and for whom a date of death **after** January 1 of the sampling year (or after the enrollment date, in the case of current-year enrollees) is recorded in CMS administrative updates or obtained from a proxy are "true" deaths, and, unlike the immortals, were alive and eligible for Medicare at the beginning of the sampling year (or as of their enrollment date, for current-year enrollees).[18] The essential difference is that the immortals are not eligible for inclusion in the MCBS since by definition they could not have incurred any health care costs in the year in which they were sampled.

---

[17] Note that members of the 2014 cohort (i.e., 2015 sampled panel members who first became eligible for Medicare during 2014) who died or lost eligibility prior to the sampling year (i.e., sometime during 2014 after becoming eligible) are not immortals and should still be sampled. These cases contribute to the 2014 Cost Supplement File. Similarly, and new beginning in 2015, members of the 2015 cohort (i.e., 2015 sampled panel members who first became eligible for Medicare during 2015) who died or lost eligibility during the sampling year (i.e., sometime during 2015 after becoming eligible) are not immortals and should still be sampled. These cases contribute to the 2015 Cost Supplement File.

[18] Data for beneficiaries in this group who were newly enrolled (i.e., enrolled during the year prior to the sampling year, or enrolled during the sampling year, beginning in 2015) are, in fact, pursued, and proxy interviews are attempted. Their data will be used to aid in imputation of their cost and use data.

For sample size determination purposes, death rates,[19] response rates, and immortal rates were computed within each age group.[20] The immortal and death rates used were an average of historical rates and Fall 2014 actual rates. The immortal rates apply to losses in the first fall interview round only. Similarly, the initial losses due to deaths in the sample selection year apply only to the first fall interview round. On the other hand, persons who completed one or more rounds of interviews but who later died in year *t* are eligible for inclusion in the Cost Supplement File covering year *t.* In other words, these later deaths do not necessarily result in a reduction in sample size in the Cost Supplement File corresponding to the year in which the beneficiary died, but do represent losses in the *subsequent* Cost Supplement Files. Thus, the "first-" and "second-year" death rates that were computed for sample design purposes are used to estimate losses in the second and third Cost Supplement Files, respectively, in which a particular panel can appear. Exhibit 3.4.1 below displays the assumed rates used in determining the sample sizes for the MCBS 2015 Panel. These rates were used in each of the Hispanic, non-Hispanic, and Puerto Rico sampling strata within age group.

---

[19] Included in the calculation of death rates is a small number of persons who lost Medicare eligibility.

[20] Note that during Fall 2014 (Round 70), a decision was made by CMS to replace any newly sampled (Incoming Panel) beneficiaries found to be incarcerated in the first interview because they would not be eligible for benefits. These numbers are quite small and are currently not significant enough to warrant inclusion in the calculation of the sample size for the annual panel.

Exhibit 3.4.1: Assumed Rates (in Percent) Used in Determining Sample Sizes for the MCBS 2015 Panel, by Age Group

Sampling Rate	Age Group (as of 7/1/2015)							
	<45	45-64	65-69	70-74	75-79	80-84	85+	Total
Estimated "immortal" rate	0.6	0.5	0.2	0.3	0.6	0.7	1.8	0.7
Estimated selection year death rate	1.8	2.2	1.6	2.2	3.2	4.7	10.9	4.1
Selection year response rate	65.2	69.7	66.6	64.1	65.0	64.9	61.4	65.0
Post-fall round death/loss rate	1.1	1.1	0.6	0.7	0.8	1.0	1.7	1.0
First year response rate	76.1	78.2	78.0	80.1	80.4	79.6	81.8	79.4
Estimated first year death rate	2.7	2.5	1.9	2.1	3.8	6.1	11.8	4.5
Second year response rate	84.9	91.2	92.8	92.0	90.9	91.8	89.8	90.9
Estimated second year death rate	3.0	3.7	1.8	2.8	3.8	7.2	13.8	5.1
Third year response rate	92.6	95.0	92.7	95.2	95.2	94.4	92.7	94.1
Year 1 Retention rate[1]	47.8	52.5	50.6	49.7	49.9	48.4	43.0	48.6
Year 2 Retention rate[2]	82.6	89.0	91.1	90.0	87.5	86.2	79.3	86.8
Year 3 Retention rate[3]	89.8	91.5	91.0	92.6	91.6	87.5	79.9	89.3

[1] The Year 1 Retention rate takes into account the immortal rate, selection year death and response rates, post fall round death/lost entitlement rate, and first year response rate. Year 1 refers to the first year after the selection year.

[2] The Year 2 Retention rate takes into account the Year 1 death rate and the Year 2 response rate. Year 2 refers to the second year after the selection year.

[3] The Year 3 Retention rate takes into account the Year 2 death rate and the Year 3 response rate. Year 3 refers to the third year after the selection year.

The response rate for the selection year used in the sample size calculations (i.e., the proportion of sampled beneficiaries, excluding deaths and immortals, who complete the initial fall interview) was assumed to be 65 percent. This assumption is based on Fall 2014 achievement and historical rates. The response rate for the first year in the survey (i.e., the proportion of persons completing the initial fall interview who provide substantially complete data for the first Cost Supplement File to which they contribute) was computed based on the average response rates for the 2009 and 2010 Panels. The response rate for the second year in the survey (i.e., the proportion of living respondents in the first Cost Supplement File who also provide substantially complete data for the second Cost Supplement File) was computed based on the average rates for the 2008 and 2009 Panels. Finally, the response rate for the third year in the survey (i.e., the proportion of living respondents in the second Cost Supplement File who also provide substantially complete data for the third Cost Supplement File) was based on the average rates for the 2007 and 2008 Panels.

The sample size projections also included adjustments to account for movement of beneficiaries from one age category to the next over the course of three years in the study. This adjustment affects primarily the age categories under 45, 85 years or older, 65 to 69, and 70 to 74 years old. As the panel ages, the oldest beneficiaries in the under 45 age category will move to the next age category and there will be no migration into the under 45 age category. There will not be any migration out of the oldest age category (85 years or older), while about 17 to 19 percent of the beneficiaries from the 80 to 84 year-old age group will move into this age group after one year. The 65 to 69 year-old age category will also be affected as the migration into this category from the 45 to 64 year-old age category will be less (about 6 to 7 percent) than the migration out of this category (about 21 to 28 percent) every year. The remaining age categories (70 to 74, 75 to 79, and 80 to 84) are not affected as much because migration in and out of these categories occurs at approximately the same rate.

The base[21] sample size target for the 2015 Panel was determined to be 8,504 beneficiaries. Exhibit 3.4.2 displays the initial base sample size planning numbers and resulting projected Survey File completed interviews in 2015 and projected Cost Supplement File completed interviews in 2016, 2017, and 2018.

The Hispanic oversample was decided upon separately; a target of 75 additional Hispanic completes per year beginning in 2015 was added to the sample. As a result, in addition to the 2015 Panel base sample of 8,504, 123 flagged Hispanic cases were added to the total sample to be selected, with a target of achieving 75 additional Hispanic completed interviews in Fall 2015, and a total of 1,500 Hispanic cases in the Survey File by 2018. Exhibit 3.4.3 below displays the projected Survey File Hispanic counts with and without oversampling.

---

[21] This excludes the Hispanic oversample, which was planned separately from but selected concurrently with the base sample. The core targeted sample size for the 2015 Panel was 8,627 beneficiaries, made up of a base sample of 8,504 beneficiaries and an oversample of 123 core Hispanic-flagged beneficiaries.

Exhibit 3.4.2: 2015 Fall Planning Sample Sizes and Resulting Projected Survey File and Cost Supplement File Estimates, by Age Group, Including Recent Enrollees and Excluding Hispanic Oversample

Age Category	Number to be Sampled for 2015 Panel: Fall 2015	Projected Starting Sample Across All Panels: Fall 2015
<45	845	1,743
45-64	513	1,343
65-69	1,304	3,252
70-74	1,337	2,759
75-79	1,468	3,101
80-84	1,475	3,085
85+	1,562	3,152
Total	8,504	18,436
Age Category	**Projected Total Survey File Completes from All Panels: 2015 Annual Estimates**	**Projected Total Cost Supplement File Completes from All Panels: 2016 Annual Estimates**
<45	1,352	857
45-64	1,083	967
65-69	2,607	1,977
70-74	2,145	1,581
75-79	2,405	1,593
80-84	2,371	1,604
85+	2,324	1,600
Total	14,287	10,179
Age Category	**Projected Total Cost Supplement File Completes from All Panels: 2017 Annual Estimates**	**Projected Total Cost Supplement File Completes from All Panels: 2018 Annual Estimates**
<45	948	1,001
45-64	996	1,000
65-69	1,946	1,900
70-74	1,836	1,900
75-79	1,815	1,900
80-84	1,819	1,900
85+	1,844	1,900
Total	11,205	11,501

Note: As described earlier in this section, recent enrollees from 2014 and 2015 are included in these estimates.

Exhibit 3.4.3: Expected Completed Interviews with Self-Reported Hispanic Beneficiaries, with and without Oversampling, by 2018

Panel	Expected Total Interviews without Over-sampling	Expected Hispanic Interviews without Over-sampling	Expected non-Hispanic Interviews without Over-sampling	Additional Hispanic Interviews as a Result of Over-sampling	Expected Total Interviews with Over-sampling	Expected Hispanic Interviews with Over-sampling	Expected non-Hispanic Interviews with Over-sampling
2015 Survey File							
2012	2,183	203	1,980	-	2,183	203	1,980
2013	2,528	235	2,293	-	2,528	235	2,293
2014	4,401	410	3,991	-	4,401	410	3,991
2015	5,175	482	4,693	75	5,250	557	4,693
Total	14,287	1,331	12,956	75	14,362	1,406	12,956
2016 Survey File							
2013	2,117	197	1,920	-	2,117	197	1,920
2014	3,553	331	3,222	-	3,553	331	3,222
2015	3,590	334	3,256	53	3,643	387	3,256
2016	5,175	482	4,693	75	5,250	557	4,693
Total	14,435	1,344	13,091	128	14,563	1,472	13,091
2017 Survey File							
2014	3,020	281	2,739	-	3,020	281	2,739
2015	2,872	267	2,605	42	2,914	309	2,605
2016	3,623	337	3,285	53	3,675	390	3,285
2017	5,175	482	4,693	75	5,250	557	4,693
Total	14,690	1,368	13,321	170	14,859	1,538	13,321
2018 Survey File							
2015	2,441	227	2,214	36	2,477	263	2,214
2016	2,898	270	2,628	42	2,940	312	2,628
2017	3,623	337	3,285	53	3,675	390	3,285
2018	5,175	482	4,693	75	5,250	557	4,693
Total	14,137	1,317	12,820	205	14,342	1,522	12,820

NOTE: Totals before and after oversampling include Puerto Rico completes; oversampled completes are only outside of Puerto Rico.

2015 Sampling Frame

As described earlier, the inclusion of current-year enrollees in the sampling frame requires additional steps to be taken in the building of the frame and sampling of beneficiaries from the frame because not all year *t* enrollees are included in the enrollment data by the spring of year *t*, when initial sampling operations occur. Instead, year *t* enrollees are added to the enrollment data in two distinct manners. First, beneficiaries who will be automatically enrolled in Medicare appear in the enrollment data up to four months prior to their automatic enrollment. Second, beneficiaries can self-enroll within a seven-month window comprised of the three months prior to their 65th birthday month, their 65th birthday month, and the three months following their 65th birthday month. Those who self-enroll typically appear in the enrollment data within a month after their enrollment in Medicare. Thus, someone enrolling in December 2015 may not appear in the enrollment data until January 2016.

In April 2015, when the first enrollment data extract was pulled to facilitate sampling for the fall round, only a portion of the current-year enrollees were included in the enrollment data. Beneficiaries who enrolled prior to April 1 of year *t* or who were automatically enrolled within four months of April (i.e., by August 1 of year *t*) were included in the enrollment data extract. However, any beneficiary who self-enrolled on or after April 1 or was automatically enrolled after August 1 of year *t* did not yet appear in the enrollment data. Thus, multiple enrollment data extracts were required to facilitate sampling of the full year *t* cohort.

Two additional enrollment data extracts are pulled each year and contribute to the year *t* sampling frame: (1) an extract in August, which contains additional self-enrollees through August 1 of year *t* and scheduled automatic enrollees through December 1 of year *t*; and (2) an extract in late September, which contains additional self-enrollees through September of year *t* and scheduled automatic enrollees through December 31 of year *t*. The September extract is scheduled for the latest date possible to facilitate sampling and fielding in year *t*; however, it leaves a slight under-coverage of any self-enrollees between October 1 and December 31 of year *t*. A final extract is pulled in mid-January of year *t+1* to identify this under-coverage and account for it in weighting adjustments.

For the 2015 MCBS Panel, the first, or initial, extract of the enrollment data, delivered in April, included:

- Beneficiaries who were first eligible for Medicare before January 1, 2014 and still alive and eligible on January 1, 2015; and
- Beneficiaries who were first eligible for Medicare between January 1, 2014 and April 1, 2015 (inclusive), or who would be automatically enrolled in Medicare between April 2, 2015 and August 1, 2015 (inclusive), regardless of vital status.

To avoid duplication across the various panels of MCBS beneficiaries, a unique and disjoint 5-percent sample of the enrollment data[22] is specified annually by CMS, and a subset (based on the eligibility and mortality selection criteria listed above, as well as other data quality checks) is specified for the MCBS for use in sampling beneficiaries for the annual panels. This is referred to as the 2015 enrollment data subsample.

CMS subset each of its enrollment data extracts as described above, keeping only beneficiaries meeting the criteria for the 2015 enrollment data subsample. These enrollment data subsample extracts are further subset to include only beneficiaries falling within the 703 selected MCBS SSUs. Exhibit 3.4.4 shows the number of beneficiaries by sampling stratum (age group by ethnicity or Puerto Rican residency) in the three 2015 enrollment data subsample extracts and the resulting 2015 sampling frame. Of the 2,872,959 beneficiaries in the combined 2015 enrollment data subsample extracts, a total of 46,568 beneficiaries fell within the selected MCBS PSUs and SSUs and were eligible for sampling in 2015.

Exhibit 3.4.4: Number of Beneficiaries in 2015 Enrollment Data Subsample Extracts (Combined) and 2015 Sampling Frame, by Stratum

Stratum	Age Group/ Ethnicity	Three Extracts Combined	2015 Sampling Frame
1	<45, Hispanic	10,850	107
2	45-64, Hispanic	34,148	412
3	65-69, Hispanic	46,315	569
4	70-74, Hispanic	46,867	568
5	75-79, Hispanic	33,599	443
6	80-84, Hispanic	22,768	280
7	85+, Hispanic	23,342	255
8	<45, non-Hispanic	85,615	1,308
9	45-64, non-Hispanic	323,480	5,252
10	65-69, non-Hispanic	715,334	10,965
11	70-74, non-Hispanic	531,723	8,246
12	75-79, non-Hispanic	379,509	5,896
13	80-84, non-Hispanic	275,315	4,385
14	85+, non-Hispanic	326,707	4,979
15	<45, Puerto Rican	474	86
16	45-64, Puerto Rican	2,765	503
17	65-69, Puerto Rican	3,981	682
18	70-74, Puerto Rican	3,455	590

---

[22] The enrollment data include over 100,000,000 beneficiaries.

Stratum	Age Group/ Ethnicity	Three Extracts Combined	2015 Sampling Frame
19	75-79, Puerto Rican	2,592	422
20	80-84, Puerto Rican	1,888	287
21	85+, Puerto Rican	2,232	333
Total		2,872,959	46,568

Using the initial 2015 enrollment data subsample extract in combination with previous annual enrollment data subsamples, the size of the total 2015 enrollment data subsample (containing all projected 2015 Medicare enrollees, through December 31, 2015) could be forecast at the time of initial sampling (June 2015). This forecast was used to determine how much of the current-year enrollee sample was expected to be selected from the first extract and how much would be expected to be drawn from future extracts, and to determine the sampling fractions for beneficiaries.

A final enrollment data subsample extract was provided in mid-January 2016 and used to fully enumerate the 2015 cohort to (a) inform undercoverage of the 2015 sampling frame, and (b) contribute to weighting adjustments to account for this undercoverage. Results of these analyses are provided in the Coverage Analysis section below.

Sample Selection for the 2015 Panel

The 2015 Panel sample was selected using the following guidelines:

1. Obtain a base sample of 8,504 beneficiaries, with targeted oversamples in the 64 and younger and 85 and over age groups; and
2. Oversample Hispanic beneficiaries (from U.S. portion of sampling frame) with a goal of interviewing 75 additional true (self-reported) Hispanic beneficiaries in 2015.

Sampling Fractions. As discussed earlier, the ethnicity flag provided by CMS and used for sampling is imperfect. Some cases flagged as Hispanic are in reality non-Hispanic, and others flagged as non-Hispanic actually self-report as Hispanic. The sampling fractions for the Hispanic and non-Hispanic strata, therefore, were jointly determined to compensate for the misclassification errors inherent in the ethnicity flag to achieve the required sample sizes of self-reported Hispanic and non-Hispanic beneficiaries.

The calculation of the sampling fractions for 2015 was accomplished in two phases. First, the sampling fractions for the U.S. (not including Puerto Rico) portion of the sample were completed at the national level within the 14 strata (seven age groups by ethnicity flag [Hispanic/non-Hispanic]). Then, the sampling fractions for the Puerto Rico portion of the sample were calculated within the seven age group strata.

Probabilities of Selection. The probabilities of selection for beneficiaries were then computed. Let f_{1a} be the national sampling fraction for the Hispanic stratum in age group a, let f_{-1a} be the national sampling fraction for the non-Hispanic stratum in age group a, and let f_{2a} be the sampling fraction for the Puerto Rican stratum in age group a. The inclusion probability for the i-th PSU is denoted by π_i and the conditional inclusion probability for the j-th SSU given the i-th PSU is $\pi_{j|i}$. Thus, the conditional probability of selection for beneficiary k in the Hispanic stratum in age group a given PSU i and SSU j is

$$\rho_{1ak|ij} = \min\left(1, \frac{f_{1a}}{\pi_i \pi_{j|i}}\right), \quad a = 1, \dots, 7,$$

and for non-Hispanic beneficiary k in the non-Hispanic stratum in age group a given PSU i and SSU j is

$$\rho_{-1ak|ij} = \min\left(1, \frac{f_{-1a}}{\pi_i \pi_{j|i}}\right), \quad a = 1, \dots, 7.$$

Because the sampling fractions for the U.S. Hispanic strata were increased to account for the Hispanic beneficiary oversample, and because the ethnicity flag was not available for current-year enrollees due to a lag in data availability, the current-year enrollees were under-selected as part of the U.S. non-Hispanic strata. To obtain a more appropriate number of current-year enrollees in the sample, the probabilities of selection for these cases[23] were doubled. As a result, the probabilities of selection for current-year enrollees are calculated as

$$\rho_{1ak|ij} = \min\left(1, \frac{f_{1a} \times 2}{\pi_i \pi_{j|i}}\right), \quad a = 1, \dots, 7,$$

and

$$\rho_{-1ak|ij} = \min\left(1, \frac{f_{-1a} \times 2}{\pi_i \pi_{j|i}}\right), \quad a = 1, \dots, 7.$$

Finally, the probabilities of selection for beneficiaries in the Puerto Rico PSUs are calculated as follows:

$$\rho_{2ak|ij} = \min\left(1, \frac{f_{2a}}{\pi_i \pi_{j|i}}\right), \quad a = 1, \dots, 7.$$

Actual sample sizes can fall short of expectations when SSUs actually contain fewer beneficiaries in the enrollment data subsample extract than what is called for by the initial national sampling

---

[23] Because of the lag in availability of the ethnicity flag for beneficiaries (it was only available for cohorts 2013 and earlier at the time of sampling of the 2015 Panel), no current-year enrollees were coded as Hispanic. Because the current-year enrollees all fell into the non-Hispanic strata, absent the doubling of the sampling rates, current-year Hispanics would have been sampled at a lower rate than if the flag had been available, resulting in an undersampling of current-year Hispanic beneficiaries. Doubling the rates is an approximate device, in the face of the missing ethnicity flag, to ensure achievement of the desired oversample.

fractions. To avoid a shortfall, the initial sampling fractions must be adjusted and the conditional probabilities $\rho_{1ak|ij}$, $\rho_{-1ak|ij}$, and $\rho_{2ak|ij}$ recomputed. Within each stratum, the cumulative sums of the probabilities of selection were formed. In an iterative process, the initial national sampling fractions were repeatedly adjusted until the cumulative sums were as close as possible to the final targeted sample sizes. Exhibit 3.4.5 displays the final sampling fractions used for calculating probabilities of selection, by stratum, for the 2015 Panel.

Exhibit 3.4.5: 2015 MCBS Panel, Final Sampling Fractions by Stratum

Stratum	Age Group/ Ethnicity	Final Sampling Fraction, in Percent
1	<45, Hispanic	0.0111
2	45-64, Hispanic	0.0020
3	65-69, Hispanic	0.0036
4	70-74, Hispanic	0.0037
5	75-79, Hispanic	0.0051
6	80-84, Hispanic	0.0079
7	85+, Hispanic	0.0104
8	<45, non-Hispanic	0.0109
9	45-64, non-Hispanic	0.0016
10	65-69, non-Hispanic	0.0016
11	70-74, non-Hispanic	0.0027
12	75-79, non-Hispanic	0.0042
13	80-84, non-Hispanic	0.0058
14	85+, non-Hispanic	0.0058
15	<45, Puerto Rican	0.0263
16	45-64, Puerto Rican	0.0030
17	65-69, Puerto Rican	0.0046
18	70-74, Puerto Rican	0.0060
19	75-79, Puerto Rican	0.0088
20	80-84, Puerto Rican	0.0123
21	85+, Puerto Rican	0.0118
Total		0.0036

Selection of the 2015 Panel. The 2015 Panel was drawn by systematic random sampling with probability proportional to the conditional probabilities of selection with an independently selected random start within each PSU. For the sample drawn from the U.S. (outside Puerto Rico), the beneficiaries were ordered within each PSU by age group, SSU (to approximate geographic serpentine sorting), ethnicity flag, and extract[24]. There were 323 beneficiaries with a conditional probability of selection equal to 1 in the first extract sample, 20 in the second extract sample, and 18 in the third extract sample. These beneficiaries were selected with certainty, given the selection of their PSUs and SSUs.

The same process was used for the sample drawn from Puerto Rico. There were no beneficiaries in Puerto Rico with a conditional probability of selection equal to 1.

Sampling Results

Exhibit 3.4.6 below shows the number of selected beneficiaries within each age group, and Exhibit 3.4.7 shows the number of selected beneficiaries within each stratum. These tables present the total number of beneficiaries in the 2015 Panel, including the Hispanic oversample.

Exhibit 3.4.6: 2015 MCBS Panel, Number of Beneficiaries Selected by Age Group

Age Group (as of December 31, 2015)	Total Selected Beneficiaries
<45	830
45-64	526
65-69	1,235
70-74	1,347
75-79	1,472
80-84	1,503
85+	1,708
Total	8,621

Exhibit 3.4.7: 2015 MCBS Panel, Number of Beneficiaries Selected by Stratum

Stratum	Age Group/Ethnicity	Total Selected Beneficiaries
1	<45, Hispanic	68
2	45-64, Hispanic	47
3	65-69, Hispanic	107

---

[24] The second extract was added to the end of the first extract, in the same sort order, and the systematic selection was continued into the range of newly enrolled beneficiaries. The same process was used for the third extract.

Stratum	Age Group/Ethnicity	Total Selected Beneficiaries
4	70-74, Hispanic	112
5	75-79, Hispanic	123
6	80-84, Hispanic	125
7	85+, Hispanic	142
8	<45, non-Hispanic	752
9	45-64, non-Hispanic	472
10	65-69, non-Hispanic	1,112
11	70-74, non-Hispanic	1,218
12	75-79, non-Hispanic	1,330
13	80-84, non-Hispanic	1,359
14	85+, non-Hispanic	1,544
15+16	<65, Puerto Rican	17
17	65-69, Puerto Rican	16
18	70-74, Puerto Rican	17
19	75-79, Puerto Rican	19
20	80-84, Puerto Rican	19
21	85+, Puerto Rican	22
Total		8,621

The number of current-year enrollees selected into the 2015 Panel (including the Hispanic oversample) is displayed in Exhibit 3.4.8 below.

Exhibit 3.4.8: 2015 MCBS Panel, Number of Current-Year Enrollees Selected by Age Group

Age Group	Initial Extract	Three Extracts Combined
<45	36	65
45-64	30	53
65-69	146	289
Total	212	407

Beneficiaries who enrolled in 2014 and were selected as part of the year t panel, as well as current-year enrollees (those who enrolled in 2015), have not been selected in any previous panel and could have incurred costs during 2015; thus, these beneficiaries will be included in the 2015 Cost Supplement File.[25] The inclusion of the current-year enrollees eliminates the need for including the

---

[25] These recent enrollees in the Cost Supplement File are first interviewed in the fall round and will have cost and use data collected from the date of the fall round interview through the end of the year. As a result, data from earlier in the year will rely on claims matching and imputation.

previous-year enrollees in the current-year Cost Supplement File moving forward, as they will now be included in the new panel each year (as current-year enrollees).

Several quality checks were performed after sample selection. These included the comparison of the weighted 2015 enrollment data subsample extract counts (combining all three extracts) with the corresponding weighted counts for the selected sample as well as the distributions of selected beneficiaries by PSU and SSU.

Coverage Analysis of the 2015 Sampling Frame

As discussed above, a final enrollment data 5-percent file extract was provided in mid-January 2016. This extract was used to fully enumerate the 2015 cohort to (a) inform undercoverage of the 2015 sampling frame, and (b) contribute to weighting adjustments to account for this undercoverage. The results of the analysis of this extract are given in this section.

Coverage Analysis. The fourth enrollment data subsample extract, along with the first three extracts, was used to fully enumerate both the 2015 enrollment data subsample and the 2015 MCBS population. In order to construct the full 2015 enrollment data subsample, all records of eligible beneficiaries enrolled through December 31, 2015, from the four extracts were combined. From that universe, the 2015 MCBS population was constructed by retaining only beneficiaries falling into the MCBS PSUs and SSUs. Including the fourth extract, which contains beneficiaries who were automatically enrolled or self-enrolled through the end of 2015, ensures that all eligible beneficiaries, particularly current-year enrollees who were not included in the first three extracts, are included in the final population. Thus, the final 2015 MCBS population includes all beneficiaries who were enrolled in Medicare in 2015 and reside in the MCBS PSUs and SSUs.

Exhibit 3.4.9 displays the full 2015 enrollment data subsample and the estimated 2015 eligible U.S. Medicare population, by stratum. This table builds on Exhibit 3.4.4, which displayed the 2015 enrollment data subsample file through the third extract.

The fourth enrollment data subsample extract is similar in size to the third extract, at approximately 19,000 beneficiaries overall. Further, the number of cases from the fourth extract falling into the MCBS PSUs and SSUs is also very similar to those in the third extract (283 in the fourth extract versus 301 in the third extract) Overall, the fourth extract accounts for 0.6 percent of the total 2015 MCBS population. While the cases included in the fourth extract consist exclusively of new enrollees, the exclusion of this extract from the frame could lead to imbalances in the representativeness of the sample. However, because the final extract accounts for such a small proportion of the overall population, it was expected to have minimal impact on the representativeness of the 2015 Panel. Any imbalance will be accounted for in adjustments made to the weights, discussed in Chapter 8.

Exhibit 3.4.9: Number of Beneficiaries in 2015 Enrollment Data Subsample and Estimated 2015 MCBS Population, by Stratum

Stratum	Age Group/ Ethnicity	Beneficiaries in Four Enrollment Data Extracts Combined	Estimated Beneficiaries in Full U.S. Medicare Population
1	<45, Hispanic	10,881	217,620
2	45-64, Hispanic	34,201	684,020
3	65-69, Hispanic	46,766	935,320
4	70-74, Hispanic	46,867	937,340
5	75-79, Hispanic	33,599	671,980
6	80-84, Hispanic	22,768	455,360
7	85+, Hispanic	23,342	466,840
8	<45, non-Hispanic	85,988	1,719,760
9	45-64, non-Hispanic	324,604	6,492,080
10	65-69, non-Hispanic	732,278	14,645,560
11	70-74, non-Hispanic	531,724	10,634,480
12	75-79, non-Hispanic	379,509	7,590,180
13	80-84, non-Hispanic	275,315	5,506,300
14	85+, non-Hispanic	326,707	6,534,140
15	<45, Puerto Rican	476	9,520
16	45-64, Puerto Rican	2,771	55,420
17	65-69, Puerto Rican	4,064	81,280
18	70-74, Puerto Rican	3,455	69,100
19	75-79, Puerto Rican	2,592	51,840
20	80-84, Puerto Rican	1,888	37,760
21	85+, Puerto Rican	2,232	44,640
Total		2,892,027	57,840,540

Exhibit 3.4.10 compares the original forecast of the full 2015 enrollment data subsample, including all cases expected to be in the enrollment data through the end of the 2015, to the actual count of beneficiaries in the combined four enrollment data subsample extracts. As described above, the forecast was used to develop sampling fractions for use in the selection of the 2015 MCBS Panel sample. The comparisons in Exhibit 3.4.10 are given by stratum and overall. The counts are quite close; the total actual overall count is only slightly higher than the forecast (2,892,027 actual versus 2,878,818 forecast beneficiaries), and the differences by stratum are very small.

Exhibit 3.4.10: Forecast Compared to Actual Beneficiaries in Full 2015 Enrollment Data Subsample, by Stratum

Stratum	Age Group/Ethnicity	Forecast[1] of Beneficiaries in Full 2015 Enrollment Data Subsample	Actual[2] Beneficiaries in Full 2015 Enrollment Data Subsample
1	<45, Hispanic	10,840	10,881
2	45-64, Hispanic	34,104	34,201
3	65-69, Hispanic	45,787	46,766
4	70-74, Hispanic	46,867	46,867
5	75-79, Hispanic	33,599	33,599
6	80-84, Hispanic	22,768	22,768
7	85+, Hispanic	23,341	23,342
8	<45, non-Hispanic	85,156	85,988
9	45-64, non-Hispanic	323,217	324,604
10	65-69, non-Hispanic	720,564	732,278
11	70-74, non-Hispanic	532,816	531,724
12	75-79, non-Hispanic	379,921	379,509
13	80-84, non-Hispanic	275,509	275,315
14	85+, non-Hispanic	326,780	326,707
15	<45, Puerto Rican	476	476
16	45-64, Puerto Rican	2,759	2,771
17	65-69, Puerto Rican	4,136	4,064
18	70-74, Puerto Rican	3,462	3,455
19	75-79, Puerto Rican	2,594	2,592
20	80-84, Puerto Rican	1,889	1,888
21	85+, Puerto Rican	2,232	2,232
Total		2,878,818	2,892,027

NOTE: The Full 2015 Enrollment Data Subsample in this table includes all current-year enrollees through December 31, 2015.

[1] Forecast was calculated at the time of sampling (June 2015).

[2] Actual counts based on all enrollment data records received for 2015, including those in final extract delivered in January 2016.

3.5 Continuing Sample (2012-2014 Panels)

Each continuing panel is fielded, along with the newly selected incoming panel, according to its rotation schedule. Panels are fielded for a total of 12 rounds, starting in the fall round of the year the panel is selected. In Winter/Summer 2015, the 2011 Panel completed its rotation schedule and was replaced by the 2015 Panel in Fall 2015. The 2012 Panel was in its 10th round of participation in Fall 2015, the 2013 Panel was in its 7th round, and the 2014 Panel was in its 4th round.

3.6 Fielded Sample Sizes by Panel and Round

During 2015, sampled beneficiaries were interviewed during two rounds.[26] There was a combined winter and summer round in 2015 and a fall round. During Winter/Summer 2015, the 2011 Panel was interviewed for its final time, and in Fall 2015, the 2015 Panel was interviewed for its first time. The fielded sample sizes,[27] by panel, for each round are given in Exhibit 3.6.1.

Exhibit 3.6.1: 2015 Fielded Sample Sizes by Round for Each Panel

Round	Panel	Fielded Sample Sizes
Winter/Summer 2015	2011	2,470
	2012	2,764
	2013	3,382
	2014	6,343
	All	14,959
Fall 2015	2012	2,470
	2013	2,918
	2014	5,147
	2015	8,621
	All	19,156

---

[26] Only two rounds of data collection in 2015 contributed to the 2015 data files. Some data were collected in early 2015 as part of Round 70, which would have contributed to the 2014 data files. Data contributing to the 2015 data files were collected between March and December of 2015, and in the early months (winter round) of 2016.

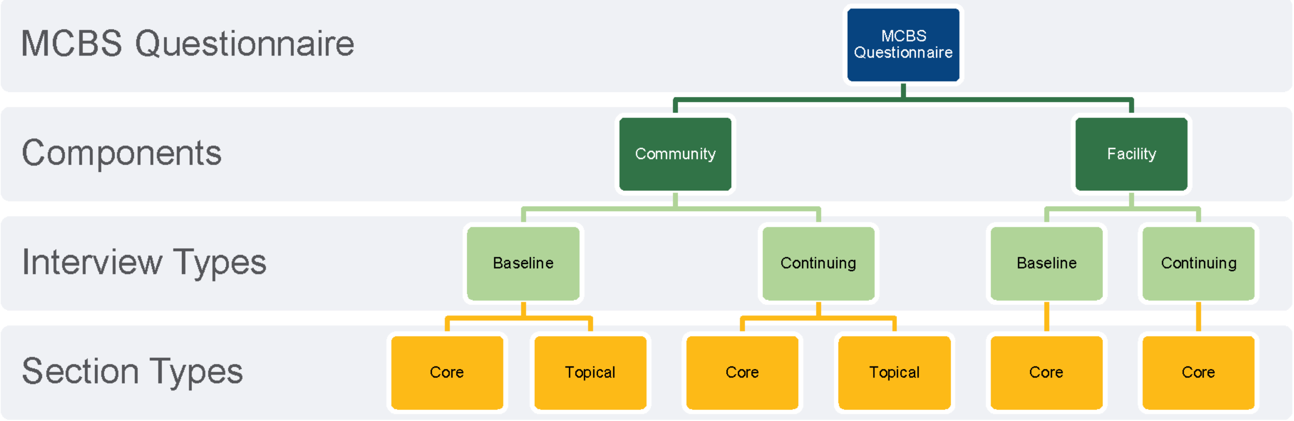
[27] Note that these are not the original sample sizes when the panel was selected (except in the case of the 2015 Panel), but the sample remaining in the round, less attrits and other sample losses, that are fielded in that round.

4. INSTRUMENT AND MATERIALS DESIGN

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 4.1 for a depiction of the MCBS Questionnaire structure.

- Community Component: Survey of beneficiaries residing in the community at the time of the interview (i.e., their residence or a household). Interview may be conducted with the beneficiary or a proxy.
- Facility Component: Survey of beneficiaries residing in facilities such as long-term care nursing homes or other institutions at the time of the interview. Interviewers do not conduct the Facility component with the beneficiary, but with staff members located at the facility (i.e., facility respondents). This is a key difference between the Community and Facility components.

Exhibit 4.1: MCBS Questionnaire Overview



Interviews are conducted in one or both components in a given data collection round, depending on the beneficiary's living situation. Procedures for these "crossover" interviews (where the beneficiary moves from one component to another) are described in Section 6.2.

Within each component, there are two types of interviews – an initial (Baseline) interview administered to new beneficiaries, and an interview administered to repeat (Continuing) beneficiaries as they progress through the study.

- 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 #2-12).

MCBS uses dependent interviewing to ensure that the flow of the interview takes into account known and previously reported information, such as beneficiary sex, health insurance coverage, health status and conditions. Dependent interviewing based on preloaded data is especially important for the design and flow of the Continuing questionnaire. This allows for a more streamlined interview by prompting the respondent for confirmation of previously-reported information, and for more complex queries to be crafted that address a beneficiary's particular situation. Section 7.2 describes the role of preloads in dependent interviewing in more detail.

Depending on the interview type and data collection season (fall, winter, or summer), the MCBS Questionnaire includes Core and Topical sections. See Exhibits 4.1.4 and 4.1.6 for tables of the 2015 Core and Topical sections.

- Core: These sections are of critical purpose and policy relevancy to the MCBS, regardless of season of administration. Core sections collect information on beneficiaries' health insurance coverage, health care utilization and costs, and operational management data such as locating information.
- Topical: These sections collect information on special interest topics. They may be fielded every round or on a seasonal basis. Specific topics may include housing characteristics, drug coverage, and knowledge about Medicare.

Data collected by the Community and Facility interviews are released to users via two primary limited data sets (LDS) – the Survey File and the Cost Supplement File. The Survey File includes data collected via Core and Topical sections related to beneficiaries' access to care, health status, and other information regarding beneficiaries' knowledge, attitudes towards, and satisfaction with their health care. This file also contains demographic data and information on all types of health insurance coverage. The Cost Supplement File delivers information collected via Core sections on the use and costs of health care services as well as information on supplementary health insurance, living arrangements, income, health status, and physical functioning.

4.1 Community Questionnaire Content

The section that follows provides an overview of the Community component of the MCBS questionnaire. The actual content administered varies based upon several factors, including the questionnaire administration season or round, the type of the interview which reflects the length of time the respondent has been in the survey, and the status as of the most recent interview.

Interview Type

As MCBS is a panel survey, the type of interview a given beneficiary is eligible for depends on his or her status in the most recent round of data collection. Interview type (also referred to in this report by its Community questionnaire variable name, INTTYPE) is a key determinant of the path followed through the Community Questionnaire. For example, the baseline interview is an abbreviated interview that includes many Core and Topical sections but does not include questionnaire sections that collect health care utilization and cost information. For the purposes of administering the Community Questionnaire, there are ten interview types, summarized in Exhibit 4.1.1 below. Because the baseline interview is always conducted in the fall, several of these interview types are applicable only in a certain season.

Exhibit 4.1.1: Community Questionnaire Interview Types

INTTYPE	Description	Seasons
C001	Standard continuing interview, meaning the most recent interview was in the community during the last round.	All
C002	New from facility, meaning the most recent interview was in a facility. No prior community interview.	All
C003	Baseline interview. First round in the sample.	Fall
C004	Standard community "holdover," meaning the last round interview was skipped. Most recent interview was in the community.	All
C005	Facility "crossover," meaning the most recent interview was in a facility. Last community interview was two rounds ago.	All
C006	Facility "crossover," meaning the most recent interview was in a facility. Last community interview was three or more rounds ago.	All
C007	Second round interview. Most recent interview was the fall baseline interview. The second round interview is the first time utilization and cost data are collected.	Winter
C008	Standard exit interview. Most recent interview was in the community during the last round. Final round in the survey.	Summer
C009	Exit interview "holdover," meaning the last round interview was skipped. Final round in the survey.	Summer
C010	Second round "holdover," meaning the winter interview was skipped. Most recent interview was the fall baseline interview. The third round interview is the first time in which utilization and cost data are collected.	Summer

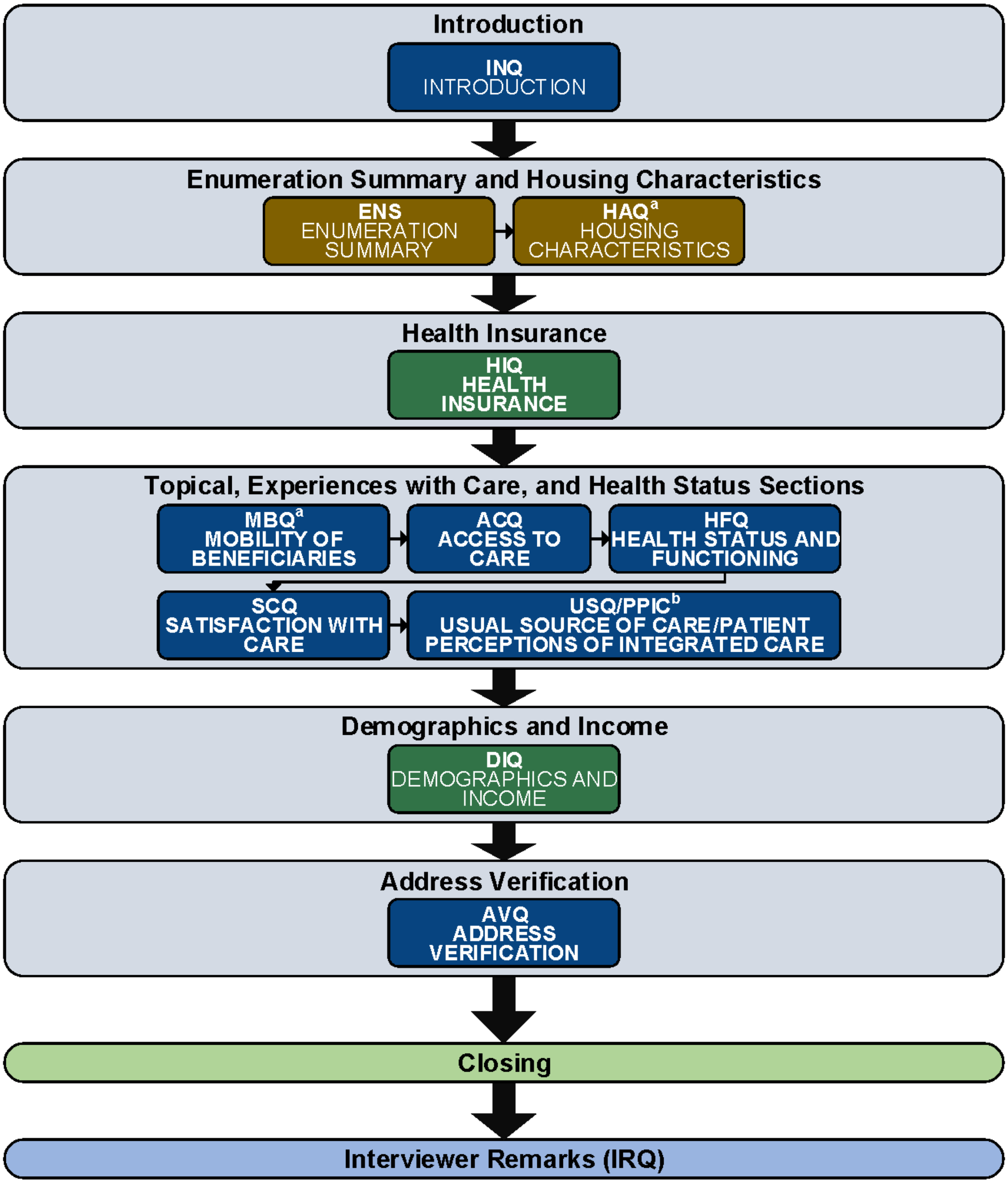
Community Questionnaire Flow

Interview type and data collection season (fall, winter, or summer) are the two main factors that determine the specific sections included in a given interview. Further factors include whether the interview is conducted with the beneficiary or with a proxy and, for proxy interviews, whether the beneficiary is living or deceased.[28] The baseline interview contains an abbreviated flow which does not include the utilization or cost sections of the questionnaire. Exhibit 4.1.2 shows the flow for the baseline interview.

---

[28] Only one final interview is conducted with proxy respondents for deceased beneficiaries. See Section 6.2 for more details on fielding procedures for deceased beneficiaries.

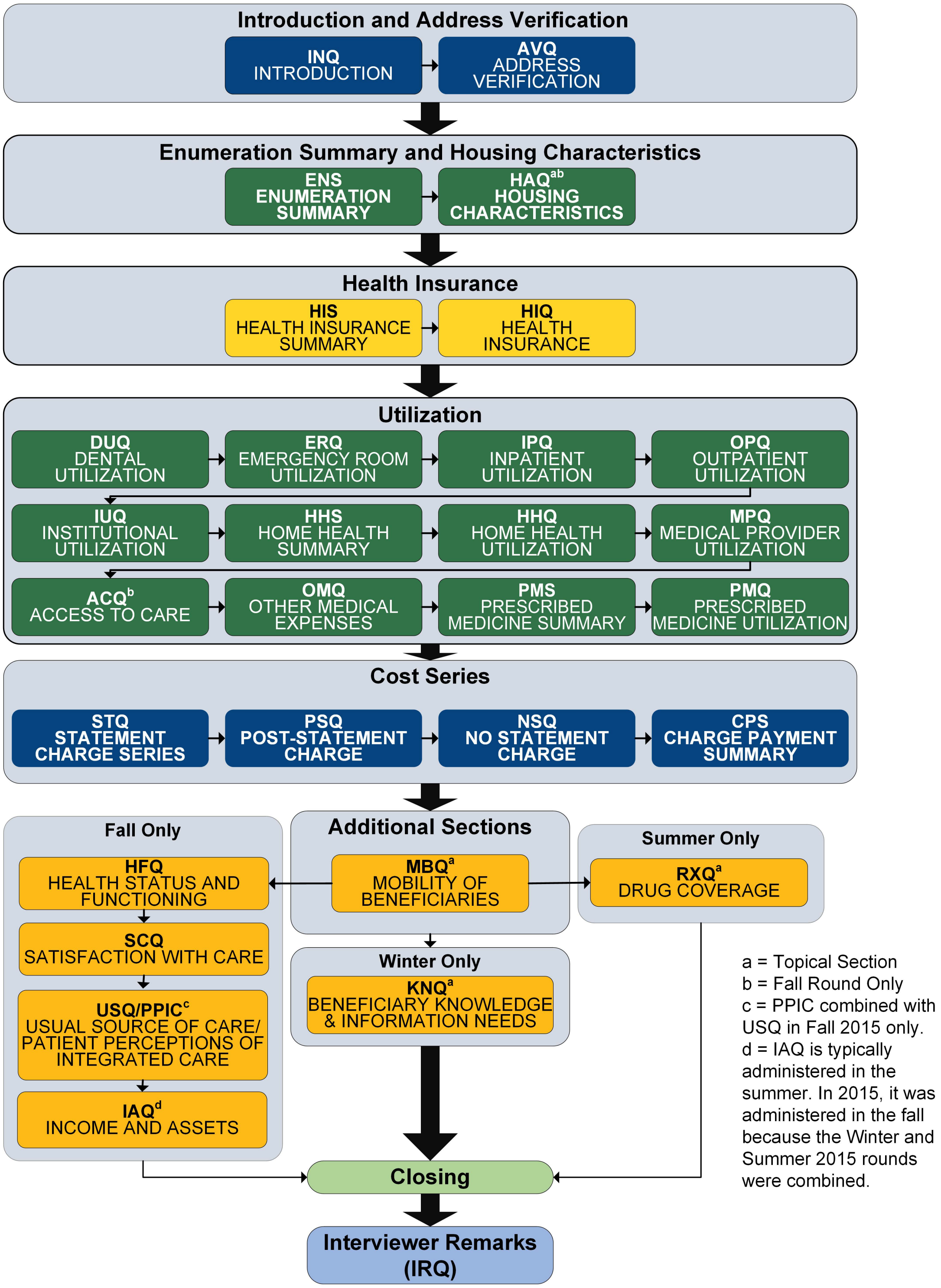
Exhibit 4.1.2: 2015 MCBS Community Questionnaire Flow for Baseline Interview



a = Topical Section
b = PPIC combined with USQ in Fall 2015 only

Exhibit 4.1.3. shows the most common Community Questionnaire flow for standard continuing community sample.

Exhibit 4.1.3: 2015 MCBS Community Questionnaire Flow for Continuing Interview



a = Topical Section
b = Fall Round Only
c = PPIC combined with USQ in Fall 2015 only.
d = IAQ is typically administered in the summer. In 2015, it was administered in the fall because the Winter and Summer 2015 rounds were combined.

Core Section Content

Core survey content is grouped into questionnaire sections that collect data central to the policy goals of the MCBS. These sections collect information related socio-demographics, health insurance coverage, health care utilization and costs, beneficiary health status and experiences with care, as well as operational and procedural data. Data from these questionnaire sections are found on the Survey File and Cost Supplement File data releases. Many of the core sections are administered each round. The following pages describe core sections of the community survey, organized by topic of information collected. Exhibit 4.1.4 lists the core sections of the community questionnaire along with the data file they inform and the seasons in which they were administered.

Socio-Demographics

Two sections in the Community Questionnaire capture key socio-demographic characteristics of the beneficiary. The Demographics and Income section is administered for each community beneficiary once during the baseline interview. Income and Assets is administered to all continuing beneficiaries once per year.

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

Income and Assets (IAQ) collects detailed information about income and assets of the beneficiary and spouse (if applicable). This section was redesigned in 2015 to mirror the income section of the National Health and Aging Trends Study (NHATS). IAQ covers beneficiary (and spouse) income from employment, Social Security, Veteran's Administration, and pensions. The respondent is also asked to indicate the value of the beneficiary's (and spouse's) assets including retirement accounts, stocks, bonds, mutual funds, savings accounts, businesses, land or rental properties, and automobiles. Also included is homeownership or rental status, and food security items. In 2015, IAQ was fielded during the fall interview, but it is normally administered during the summer round interview. Traditionally the Income and Assets section is asked in the next summer round to collect income and asset information about the previous calendar year. The Income section asked in the Fall of 2015 asked about income and assets for the 2014 calendar year.

Exhibit 4.1.4: 2015 MCBS Community Core Sections by Data File and Administration Schedule

Topic	Section Name	Survey File	Cost Supplement File	Administration Schedule
Socio-Demographics	DIQ - Demographics/Income	X		Fall, Baseline Interview
	IAQ - Income and Assets*	X		Summer
Health Insurance	HIS - Health Insurance Summary**	X		All seasons
	HIQ - Health Insurance	X		All seasons
Utilization	DUQ - Dental Utilization		X	All seasons
	ERQ - Emergency Room Utilization		X	All seasons
	IPQ - Inpatient Hospital Utilization		X	All seasons
	OPQ - Outpatient Hospital Utilization		X	All seasons
	IUQ - Institutional Utilization		X	All seasons
	HHS - Home Health Summary**		X	All seasons
Utilization	HHQ - Home Health Utilization		X	All seasons
	MPQ - Medical Provider Utilization		X	All seasons
	OMQ - Other Medical Expenses Utilization		X	All seasons
	PMS - Prescribed Medicine Summary**		X	All seasons
	PMQ - Prescribed Medicine Utilization		X	All seasons
Cost	STQ - Statement Charge Series		X	All seasons
	PSQ - Post-Statement Charge		X	All seasons
	NSQ - No Statement Charge		X	All seasons
	CPS - Charge Payment Summary**		X	All seasons
Experiences with Care	ACQ - Access to Care	X		Fall
	SCQ - Satisfaction with Care	X		Fall
	USQ/PPIC - Usual Source of Care/ Patient Perceptions of Integrated Care	X		Fall
Health Status	HFQ - Health Status and Functioning	X		Fall

SOURCE: 2015 MCBS Community Questionnaire

NOTE: 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), Address Verification (AVQ), Enumeration (ENS), Closing (CLQ), and Interview Remarks (IRQ)).

*In 2015 only, due to the combined Winter/Summer Rounds, the IAQ was fielded in the fall to Continuing respondents only. IAQ is normally fielded in summer rounds. The IAQ collects income and asset information about the previous calendar year. The 2015 IAQ collected income and asset information for the 2014 calendar year.

**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.

Health Insurance

Two sections of the community questionnaire capture health insurance information.

Health Insurance Summary (HIS) reviews information about health insurance plans that the beneficiary had at the time of the last interview. Plans reported in the prior round may be deleted, or edited in this section, if the respondent indicates they are not correct for the previous reference period. Additional plans may be added if they are active during the prior round reference period but are not reported at that time. HIS is administered for all beneficiaries that had an interview in the prior round. Therefore, newly sampled beneficiaries in the baseline interview, beneficiaries in their final interview, and those that skipped the most recent round or crossed over from the facility questionnaire do not receive this section.

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

Utilization

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

All utilization sections are administered in all continuing interviews; these sections are not part of the incoming panel’s baseline interviews. Additional detail is provided on each of the four types of health care utilization collected by the community survey below.

Provider Service Visits

The utilization sections collecting provider service dates are as follows.

Dental Utilization (DUQ) collects information about dental visits during the reference period. DUQ collects the name and type of dental providers, dates of visits, services performed, and medicines prescribed during the visits.

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

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

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

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

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

Home Health Care Visits

A second type of health care utilization captured by the community survey are home health care visits. For continuing beneficiaries that reported home health events during the prior round, **Home Health Summary (HHS)** reviews those providers and confirms whether the same providers were visited during the current round. These visits are recorded not by date, but by the number of visits. In addition, the length of visits and services performed are recorded. **Home Health Utilization (HHQ)** then collects information about home health provider visits, both professional and non-professional, during the reference period. HHQ collects names and types of home health providers, number and length of visits, and services performed during visits.

Other Medical Expenses

The community survey also records other medical expenses. These expenses are reported using a slightly different reporting structure within the questionnaire.

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

Prescribed Medicines

For continuing interviews with at least one medicine reported in the prior round, the **Prescribed Medicine Summary (PMS)** presents the list of medicines reported during the prior round interview and asks the respondent to verify that those medicines are correct as of the date of the prior round interview. This allows the interviewer to add, delete, or edit medicines applicable to the prior round reference period.

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

Cost Series

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

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

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

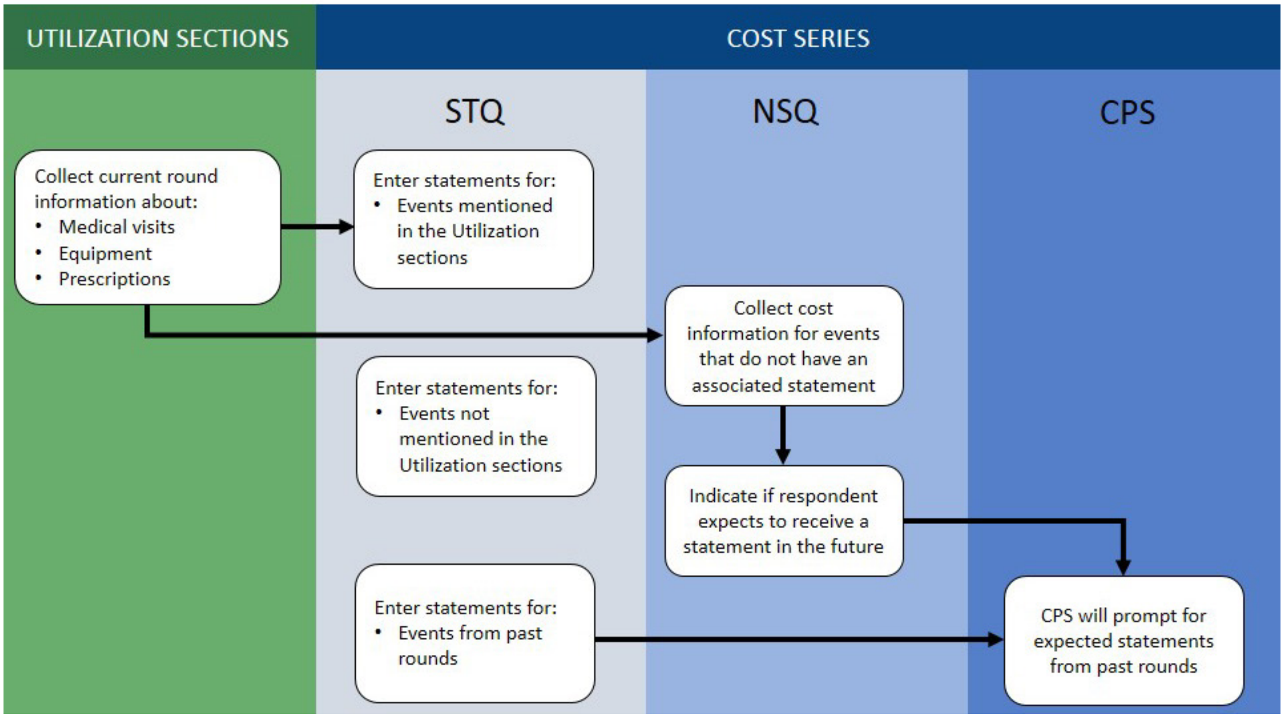
The **Post-Statement section (PSQ)** facilitates cost data collection for rental items that span multiple rounds of interviews (such as a long term wheelchair rental) and for which cost data has not yet been reported.

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

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

The flow of sections and questions within the Cost series varies depending on data collected in the current round (e.g., whether the beneficiary had a health insurance statement for a visit reported in the current round) and data collected in prior rounds (i.e., whether there was outstanding cost information reported from a prior round). Exhibit 4.1.5 illustrates how paths through these sections may vary depending on health care utilization and cost information collected in the current and previous rounds.

Exhibit 4.1.5: Utilization and Cost Section Flow



Health Status

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

Experiences with Care

Three sections cover the beneficiary's experience with care in various medical settings. All three Experiences with Care sections are part of the fall round interview for Incoming Panel and continuing respondents.

Access to Care (ACQ) focuses on the beneficiary's experience with particular types of medical encounters (hospital emergency room, hospital clinic or outpatient department, long-term care facility, or medical doctor visits) during the reference period. If the beneficiary had one or more of a particular type of medical encounter, additional items collect information about services received and waiting times associated with the most recent encounter.

Satisfaction with Care (SCQ) collects the respondent's opinions about the health care that the beneficiary had received. The questions refer to medical care received from all medical providers, including both doctors and hospitals.

In Fall 2015, the **Usual Source of Care (USQ)** section was augmented with items from the **Patient Perceptions of Integrated Care (PPIC)** module[29] to create a new section (USQ/PPIC) that obtained specific information about the usual source of health care for the beneficiary as well as any specialists seen during the reference period. The PPIC items focused on issues relating to the integration of care that patients received across providers and specialists.

Operational and Procedural

These sections help guide the interviewer through the interview, providing scripts for introducing and ending the interview. They also facilitate collection of address and household information to augment sample information for the purposes of locating respondents for follow-up interviews. Data collected in these sections are not included in the Survey or Cost Supplement data files.

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

Address Verification (AVQ), Closing (CLQ) and Exit (EXQ) sections obtain contact information necessary to locate the respondent for future rounds of interviewing. AVQ collects the beneficiary's contact information (address, phone number, alternate address, etc.), and CLQ collects contact information for the proxy (if applicable), two additional individuals who may serve as proxies in the future, and two additional contact persons who do not live in the beneficiary's household. AVQ and CLQ are administered in all rounds except the final exit interview. EXQ is administered in place of CLQ during the exit interview and contains additional scripts to thank the respondent for participation over the four years of the survey.

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

The **Interviewer Remarks Questionnaire (IRQ)** captures additional metadata about the interview, as recorded by the interviewer. This includes the length of the interview, assistance the respondent may have received, perceived reliability of the information provided during the interview, and comments the interviewer had about the interviewing situation. IRQ is administered

---

[29] Singer et al., "Development and Preliminary Validation of the Patient Perceptions of Integrated Care Survey," Medical Care Research and Review 70, no. 2.

after every interview, but is generally completed after leaving the respondent's home, as none of the questions are directed to the respondent.

Topical Section Content

In addition to the core content, there are several topical questionnaire sections that capture data on a variety of key topics that are of interest and relation to the beneficiary's housing characteristics, health behaviors, knowledge about Medicare, and health-related decision making. All data from the topical sections are included in the Survey File data release. Each topical section is described below, organized by information collected. Exhibit 4.1.6 lists the topical sections and administration schedule.

Exhibit 4.1.6: 2015 MCBS Community Topical Sections by Data File and Administration Schedule

Section Group	Abbr.	Section Name	Traditional Season	Administrative Season
Housing Characteristics	HAQ	Housing Characteristics	Fall only	Fall 2015
Health Behaviors	MBQ	Mobility of Beneficiaries	All seasons	Winter/Summer 2015, Fall 2015
	PVQ	Preventive Care	All seasons	Winter 2016
Knowledge and Decision Making	KNQ	Beneficiary Knowledge and Information Needs	Winter only	Winter/Summer 2015
	PAQ	Patient Activation	Summer only	Summer 2016
	RXQ	Drug Coverage	Summer only	Winter/Summer 2015

Housing Characteristics

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

Health Behaviors

Two questionnaire sections record additional information about health behaviors, specifically mobility and preventive care.

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

The **Preventive Care (PVQ)** section was added to the survey starting in Winter 2016. This topical section is part of the 2015 Survey File (see Chapter 7.3 Survey Data File). The Winter 2016 items within PVQ focused on the influenza vaccine. Prior to 2016, these items were administered as part of HFQ in the fall, but were moved to PVQ to allow for more accurate reporting, as flu vaccines are generally given in the fall and early winter.

Knowledge and Decision-Making

Respondent knowledge of Medicare and health-related decision making is captured in three topical sections.

The **Beneficiary Knowledge and Information Needs (KNQ)** section is administered in the winter round. These items measure the respondent's self-reported understanding of Medicare and common sources of information about health care and Medicare.

The **Patient Activation (PAQ)** section is administered during the summer round, and only when the beneficiary is the respondent. It covers items such as the beneficiary's interaction with health care providers and ability to find and understand information from those providers.

The **Drug Coverage (RXQ)** section is a summer round section that focuses on the Medicare Prescription Drug benefit, including respondent knowledge of the benefit, and opinions of the beneficiary's drug coverage, whether through a Medicare Prescription Drug Plan, a Medicare Advantage plan with prescription drug coverage, or a private insurance plan that covers prescription drugs.

Changes to the Community Questionnaire for 2015

Questionnaire changes implemented for 2015 fell into one of two categories: (1) one-time 2015 updates to questionnaire flow to accommodate an altered data collection schedule and (2) revisions to specific questionnaire items and sections. Questionnaire updates made to accommodate one-time changes to the data collection schedule in 2015 involved temporary revisions to the questionnaire flow by interview type and season. Updates to specific questionnaire items and sections generally included long-term revisions to questionnaire programming logic, addition of new questionnaire items, and updates to question text and response options.

Summary of Community Questionnaire Flow Updates Made for 2015

As a result of the combined 2015 winter and summer data collection rounds, several sections were administered in a different season than typically administered. As indicated in Exhibit 4.1.7, the Beneficiary Knowledge and Information Needs section (KNQ), usually asked in the winter was asked in the combined Winter/Summer 2015 round. Drug Coverage (RXQ), usually asked in the summer, was included in Winter/Summer 2015; Income and Assets section (IAQ), usually asked in the summer, was shifted to Fall 2015; and Patient Activation (PAQ), usually asked in the summer, was dropped in 2015[30] due to overlap with some of the new content added to the revised Usual Source of Care that included new Patient Perceptions of Integrated Care (USQ-PPIC) measures fielded in Fall 2015. These changes to the season of interview administration have minimal effect on the content of the final data files, which are annualized.

Furthermore, in 2015, slightly different rules applied for the exiting panel. The final (12th) interview is generally abbreviated and does not entail any cost or utilization sections. However, in Winter/Summer 2015, the exiting panel's final interview was the 11th interview, which is similar to a standard continuing interview. These respondents received the same interview flow as would a standard continuing case with updated scripts during the Exit section of the questionnaire to indicate that this would be the final interview.

Exhibit 4.1.7: 2015 MCBS Topical Community Questionnaire Section Changes

Traditional Season	2015 Modified Season/Round	Abbr.	Description	Survey File
All seasons	Winter/Summer, Fall	MBQ	Mobility of Beneficiaries	X
Fall only	Fall	HAQ	Housing Characteristics	X
Winter only	Winter/Summer	KNQ	Beneficiary Knowledge and Information Needs	X
Summer only	N/A	PAQ	Patient Activation	X
Summer only	Winter/Summer	RXQ	Drug Coverage	X

Summary of Item- and Section-Level Questionnaire Revisions

Charge Payment Summary (CPS)

The scope of CPS was slightly reduced for the 2015 data collection year. In prior years, there were eight distinct reasons why a cost record could be flagged for follow up in CPS. The majority of costs fall into three categories, so in 2015, only costs with those three most common CPS reasons were flagged for follow up.

---

[30] Although PAQ was not fielded in 2015, data from the PAQ administered in 2016 are available in the 2015 data files.

Demographics and Income (DIQ)

In Fall 2015, the order of three Limited English Proficiency (LEP) measures in DIQ was revised such that the measure of Primary Language was asked first in the LEP series for all Incoming Panel respondents. The purpose of this change was to align LEP items in the MCBS Community questionnaire specifications with guidance from the Department of Health and Human Services on data collection standards for primary language. In Fall 2015, the income question asked in the DIQ was revised to bring the categories closer in line to the new Income and Assets measures derived from the National Health and Aging Trends Study (NHATS).

Dental Utilization (DUQ)

Two follow-up items were added to this section for respondents reporting no dental utilization. These items captured whether the beneficiary needed dental care but could not get it, and if so, the reasons why the beneficiary could not get dental care. For respondents reporting a dental event, items in the DUQ were modified and added to capture the type of dental provider visited and the type of procedure received.

Health Status and Functioning (HFQ)

In keeping with Department of Health and Human Services guidance for data collection standards, six questions that assess disability were included. These items were added to the beginning of HFQ and collect data about vision and hearing problems, difficulty concentrating and making decisions, difficulty with activities of daily living (e.g., mobility, bathing and dressing,) and instrumental activities of daily living (e.g., difficulty running errands.)

Income and Assets (IAQ)

In Fall 2015, the IAQ was substantially revised to align the questions and approach to collecting income and assets with other surveys, such as the National Health and Aging Trends Study (NHATS).[31] In addition to replacing IAQ items with selected sections from the NHATS, six additional questions from the U.S. Department of Agriculture Economic Research Service were added to this section to provide a measure of food security.

Beneficiary Knowledge and Information Needs (KNQ)

In Winter 2016, three items were added to this section to identify some of the reasons beneficiaries find it difficult to compare health insurance plans and make plan choices.

---

[31] Montaquila, J, VA Freedman, B Spillman, and JD Kasper. "National Health and Aging Trends Study Development of Round 1 Survey Weights." *NHATS Technical Paper* 2 (2012).

Mobility of Beneficiaries (MBQ)

Response options were added to two items regarding limitations on beneficiary mobility and driving. The new response options allow interviewers to specify that the beneficiary does not drive.

Usual Source of Care – Patient Perceptions of Integrated Care (USQ-PPIC)

Items from the Patient Perceptions of Integrated Care (PPIC) instrument were integrated into the MCBS Community questionnaire as a combined module with the current Usual Source of Care (USQ) section. The PPIC asks about issues related to the integration of care that patients receive across providers. By incorporating the PPIC into the USQ, all questions pertaining to a beneficiary's usual source of care were kept together.

4.2 Facility Instrument Content

The following section provides an overview of the content of the Facility component of the MCBS questionnaire. The content of the Facility Instrument varies based upon several factors, including the season of data collection, the type of interview (which reflects the length of time the beneficiary has been in the facility), and the component of the most recent interview.

Interview Type

Similar to the Community Questionnaire, the Facility Instrument uses interview types as a key determinant of which questionnaire sections to administer during a facility interview.

The MCBS uses five categories, known as sample types, to describe MCBS beneficiaries who reside in a facility, summarized in Exhibit 4.2.1.

Exhibit 4.2.1: Facility Instrument Interview Types

INTTYPE	Description	Season
CFR	Continuing Facility Resident. Beneficiary whose previous round interview was a facility interview and who currently resides at the same facility.	Any
CFC	Community-Facility-Crossover. Beneficiary who interviewed in the community previously and has now moved to a long-term care facility.	Any
FFC	Facility-Facility-Crossover. Beneficiary who was previously interviewed in a long-term care facility and has now moved to a different facility.	Any

INTTYPE	Description	Season
FCF	Facility-Community-Facility Crossover. Beneficiary whose last interview was in the community and for whom a facility interview has been conducted in a previous round, and who has been admitted to a new facility or readmitted to a facility where the beneficiary had a previous stay. This sample type is rarely encountered.	Any
IPR	Incoming Panel Respondent. Beneficiary who was just added to the MCBS sample (fall round only) and currently resides in a facility.	Fall

NOTE: Interview type (INTTYPE) is also referred to as Sample Type in the Facility Instrument section specifications.

Facility Screener

The Facility screener is administered to a facility staff member when a beneficiary moves to a new facility setting. The Facility screener confirms whether the beneficiary is currently living at the facility (or lived at the facility at some point during the reference period) and determines whether the facility is a public or private residence.

Facility Instrument Flow

The Facility Instrument collects similar data to the Community questionnaire. However, the Facility Instrument is administered to facility staff and not to the beneficiary; that is, the beneficiary does not answer questions during a Facility interview – instead, facility administrators and staff answer questions on behalf of the beneficiary.

Just like the Community questionnaire, the sections administered in a given facility interview vary by interview type and data collection season (fall, winter, or summer). The baseline interview administered to Incoming Panel Respondents contains an abbreviated flow which does not include the utilization or cost sections of the questionnaire. Exhibit 4.2.2 shows the flow for the baseline interview.

Exhibit 4.2.2: 2015 MCBS Facility Instrument Flow for Baseline Interview

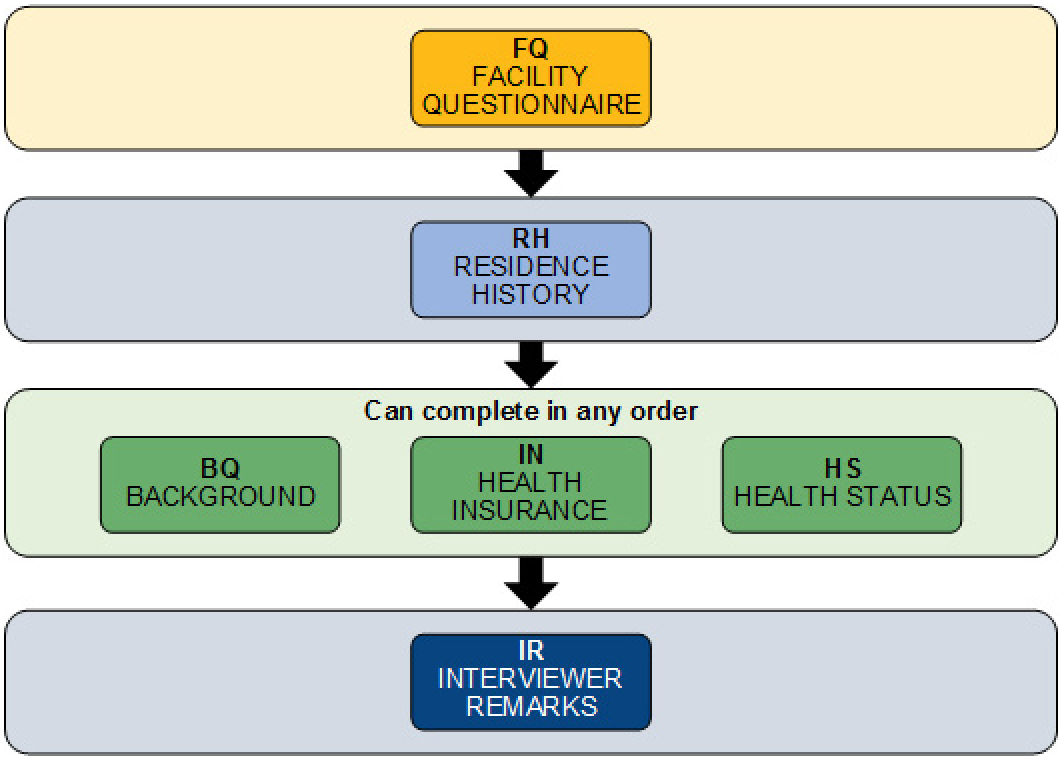
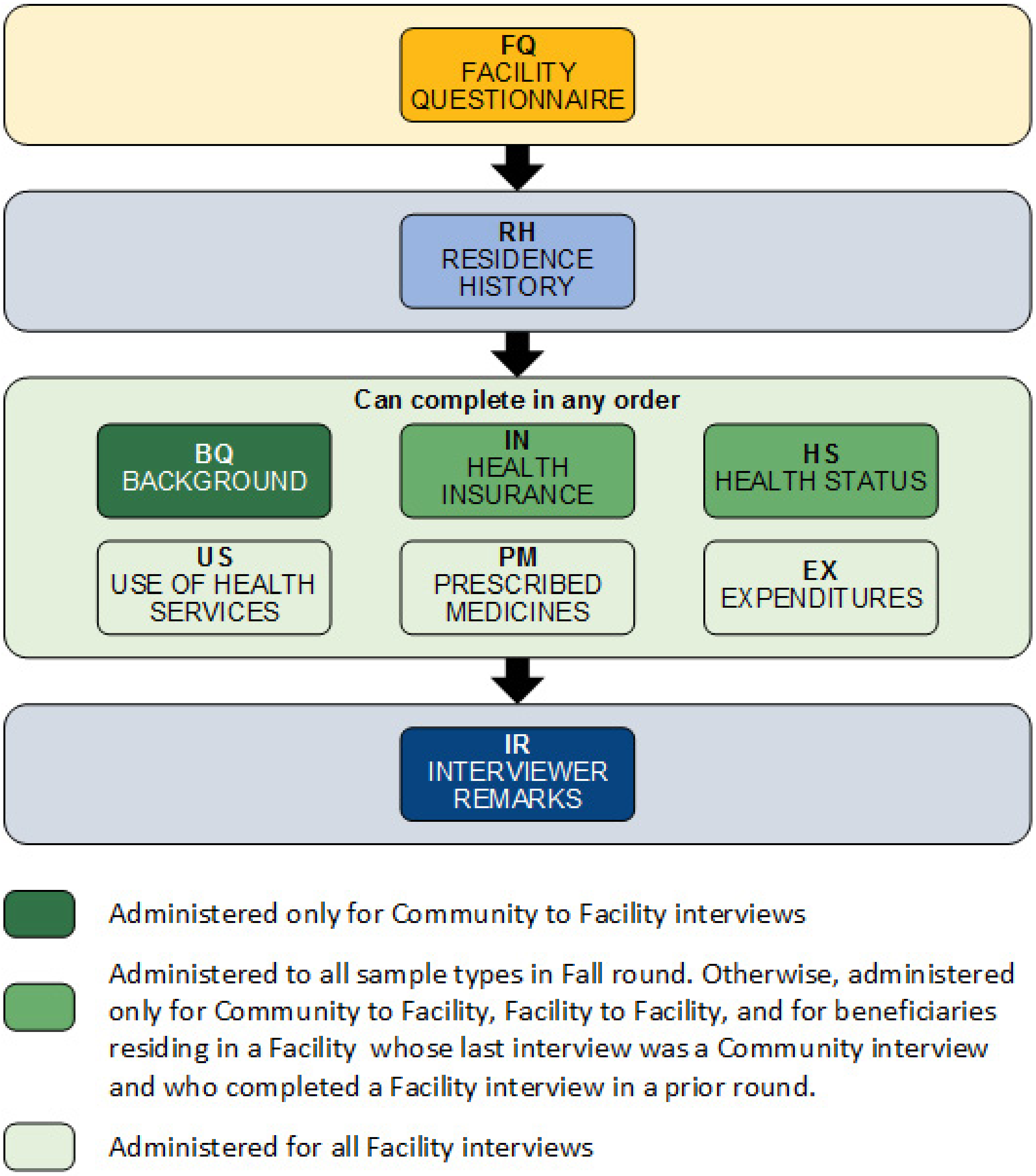


Exhibit 4.2.3 shows the flow for the continuing and crossover interview types.

Because the Facility Instrument is administered to facility staff and not directly to the beneficiary, the Facility Instrument is designed to have a modular, flexible flow. The interviewer first completes the Facility Questionnaire (FQ) section. Next, the interviewer administers the Residence History (RH) section. The remaining sections may be completed in any. Interviewers are instructed to conduct the sections in the order most suitable to the facility structure and the availability of facility staff. For example, the interviewer may conduct three sections with the head nurse and then visit the billing office to complete the remaining sections. Interviewers complete the Interviewer Remarks (IR) section at the end of the interview.

Exhibit 4.2.3: 2015 MCBS Facility Instrument Flow for Continuing and Crossover Interviews



🟩 Administered only for Community to Facility interviews

🟩 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.

☐ Administered for all Facility interviews

Core Section Content

The Facility Instrument consists of only core sections with no topical content. The following pages describe core sections of the Facility Instrument, organized by topic of information collected. Exhibit 4.2.4 shows the core sections of the Facility Instrument along with the data file they inform and the seasons in which they are administered.

Exhibit 4.2.4: Facility Core Sections by Data File and Administration Schedule

Topic	Section Name	Survey File	Cost Supplement File	Season of Administration
Facility Characteristics	FQ Facility Questionnaire	X		All seasons
Socio-Demographics	RH Residence History	X		All seasons
	BQ Background	X		Fall*
Health Insurance	IN Health Insurance	X		Fall**
Utilization	US Use of Health Services		X	All seasons
	PM Prescribed Medicines		X	All seasons
Cost	EX Expenditures		X	All seasons
Health Status	HS Health Status	X		Fall**

SOURCE: 2015 MCBS Facility Instrument

NOTE: 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.

**The IN and HS sections are also administered to Community-to-Facility and Facility-to-Facility crossover cases each season.

Facility Characteristics

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

Socio-Demographics

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

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

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

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

Health Insurance

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

Utilization

The Utilization sections collect data on the beneficiary's use of health care and prescribed medicines. These sections are administered to all sample types except for the Incoming Panel sample.

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

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

The **Prescribed Medicines (PM)** section collects data on the beneficiary's monthly use of prescribed drugs while in the facility. The data collected includes the medicine name, form, strength, dosage, how often it is administered, and the total times per month that the medicine is taken. The data for this section may be obtained by abstracting from the Medication Administration Record (MAR), which is the report that serves as a legal record of the drugs administered to a patient at a facility by a health care professional. MARs are commonly referred to as drug charts. Unlike other parts of the questionnaire, the PM section collects data by month, asking for each month separately.

In the Facility Instrument, prescribed medicines are defined as medicines ordered by a physician through a written or verbal order for a pharmacist to fill. In the long-term care setting, doctors often order medications that do not require a prescription in the community setting, such as vitamins. This is a key difference between the Facility and Community Instruments. As a result, medications that are not considered prescribed medicines in the Community Instrument are collected as prescribed medicines in the Facility Instrument.

Cost

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

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

Health Status

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

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

The HS questionnaire is designed to mirror the flow and wording of the MDS items; it contains a subset of the MDS items. In addition, the HS section contains some questions that are not found on the MDS. Interviewers ask these questions of someone knowledgeable about the beneficiary's care or find the information in the medical chart.

Operational and Procedural

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

Missing Data Sections

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

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

The missing data sections are:

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

Changes to the Facility Questionnaire for 2015

There were no substantive changes to the Facility Instrument for 2015. However, the reference period for cases in their second interview was modified to accommodate the extension of the Fall 2014 interview period. The reference period displayed for these cases during their Winter/Summer 2015 interview was modified to allow for collection of utilization and cost data from January 2015 through March 2015.

4.3 CAPI and Case Management System Programming and Testing

MCBS interviewers receive project laptops with computer assisted personal interviewing (CAPI) software and an electronic case management system to facilitate data collection activities and questionnaire administration. Interviewers conduct the MCBS interviews using the CAPI software on the laptops and organize their cases and workload using the case management system. This section describes the CAPI and case management systems.

Community Instrument

The MCBS Community Instrument used in 2015 was programmed in UNICOM® Intelligence™ Data Collection Suite (formerly IBM® SPSS® Data Collection or mrInterview). The software allows for full control of interviewer routing through the complex questionnaire. It uses built-in data quality measures such as range and logic checks, dynamic text fills, and respondent exit and re-entry management. Several lookup tools are also included within the questionnaire to allow for more effective identification of some types of health insurance plans (Medicare Advantage and

Prescription Drug plans), medical providers, and prescribed medicines. Throughout the questionnaire, specially formatted grid screens allow interviewers to easily reference providers, health care events, and medicines added in the current round, as well as those added in prior rounds (and preloaded into the questionnaire). In addition, screen-by-screen help text is available to assist interviewers with definitions and additional instruction.

Facility Screener and Instrument

The MCBS Facility Instrument is programmed in Blaise® interview software. Unlike the Community Instrument, the Facility Instrument is modular, meaning the software allows the interviewer to select sections based on the interviewing situation, rather than on a set order (with some restrictions, see Section 4.2 for more information). The Facility interview includes built-in data quality checks such as range and logic checks, dynamic text fills, respondent exit and re-entry, and a prescribed medicine lookup utility. The Facility Instrument also features a facility stay history timeline.

The Facility Screener is a separate instrument programmed in UNICOM Intelligence Data Collection Suite. This module allows for basic information about a facility to be recorded electronically and transferred to an interviewer certified to complete the facility interview.[32]

Case Management System

The case management system facilitates management of interviewer case assignments and questionnaire administration. It is a web-based application that provides interviewers and other project staff with a consistent way to access, update, and organize case information (e.g., contact names, addresses, telephone numbers, date and location of the last interview, and optimal contact time). The system includes a portal-based case management view and a laptop-based interviewing module. Field managers and other project staff use the management portal to monitor interviewer workload and productivity. Interviewers use the laptop-based module to view their MCBS case assignments, record attempts to locate and contact respondents, update respondents' personal contact information, schedule appointments, and record case status information. The case management system is the gateway for interviewers to access the Community and Facility Instruments, as well as the Facility Screener. Interviews are conducted off-line and case management and survey data are synchronized between the laptop database and the central office servers over a secure, encrypted internet connection.

Paradata elements captured within the case management system include contact level information, mode of contact attempt, source of contact information referenced (phone, address, email, etc.), and the result of the contact attempt. The case management system integrates questionnaire and

---

[32] Not all interviewers may complete Facility interviews – additional training and certification is required beyond the standard Community interview training.

case management data both within and across rounds allowing interviewers to identify the best or most recent telephone numbers and locations for expedited contacting.

In Fall 2015, an Automated Crossover Process (ACP) was implemented to automatically transfer cases from the Community component to the Facility component. The ACP creates case management updates and questionnaire preloads for these cases through a set of stored procedures, allowing interviewers to conduct an interview with the facility as quickly as one day after they located and screened the facility.

4.4 Letters and Other Respondent Materials

A series of materials and other resources provide respondents with information about the MCBS and request their cooperation and participation in the survey. Respondents receive letters by mail and additional materials from interviewers during their visits. In addition, a MCBS respondent website, a project toll-free number, and project email address are available for respondent communication.

Respondent materials include a variety of standard letters, such as advance letters mailed prior to the baseline interviews, refusal conversion letters sent to convey the importance of participation, and a community authority letter sent to communicate survey legitimacy of the survey to entities such as police departments. Materials are tailored to whether respondents reside in the community or in facilities. In addition to the standard letter mailings, a set of contacting and refusal conversion letters are used to address common contacting problems and respondent concerns about participating in the study.

Other materials include brochures about the survey, NORC, and the income and assets questionnaire section; greeting cards; a frequently asked questions (FAQ) document, and an annual newsletter. The brochures, cards, FAQs may be used by interviewers or managers at their discretion to assist in gaining cooperation. The MCBS respondent newsletter was provided to all beneficiaries residing in the community during the Winter/Spring 2015 interview. Finally, to assist in recording health care visits, purchases, and costs, all respondents to the Community interview in the fall round receive a planner booklet with a calendar that allows them to record health care and other appointments and costs.

5. INTERVIEWER RECRUITMENT AND TRAINING

5.1 Interviewer Recruitment and Staffing

A professional interviewer staff is required to complete in-person interviews throughout the year. In 2015, most MCBS interviewers were experienced and had conducted MCBS interviews for at least a year or more. Some new-to-MCBS interviewers were recruited to replace those who had left the project; hiring is targeted based on local staffing needs and MCBS-specific skill requirements. The set of preferred skills included experience with financial data and complex surveys; language skills; working with individuals who have hearing, visual, or cognitive challenges; and experience interviewing those with disabilities and the elderly. Approximately 260 interviewers worked on the MCBS in 2015.

5.2 Interviewer Training Programs for 2015

Interviewer Training Goals and Objectives

The 2015 MCBS Training Program included in-person and remote trainings, customized by level of interviewer experience (new-to-MCBS or MCBS-experienced), interview component (Community or Facility), sample type (Incoming Panel or Continuing), and season-specific requirements (new questionnaire modules or data collection protocols). The program was structured to expose all field staff to the same training content, ensuring that the performance of data collection responsibilities was standardized, methodical, and measurable.

Two in-person trainings targeted new-to-MCBS interviewers in advance of Fall 2015 data collection, while remote trainings targeted MCBS-experienced interviewers in advance of each round of data collection. In-person trainings educated new-to-MCBS interviewers on the project's background and purpose, preparations for gaining cooperation under various circumstances, and proper administration of the MCBS questionnaires to both incoming panel and continuing respondents. Remote trainings reinforced key interviewing skills and educated interviewers on round specific questionnaire sections and protocols.

In-Person Interviewer Training

New MCBS interviewers attended two in-person training sessions in 2015: the first to cover the content of the baseline interview and the second to cover the continuing interview. The 2015 in-person training sessions for new MCBS interviewers were modular in structure; each module targeted training objectives by incorporating multiple formats and media (including video, group activity, self-study, and active demonstration) to deliver training points. Practice (or "mock") interview scenarios were used to demonstrate CAPI functionality and questionnaire flow.

Baseline

The first training session introduces new interviewers to the MCBS and focuses on the Baseline interview, including fielding strategies, gaining cooperation and CAPI questionnaire administration. This training includes home study preparation and classroom learning during which trainers observe interviewers and provide them with feedback. Interviewers are required to complete a certification exercise before they may conduct interviews.

Continuing

The second training session focuses on the Continuing interview and highlights additional protocols for navigating the CAPI questionnaire. This training focuses on in-depth coverage of how to work with respondents' health care and medical documentation to obtain cost data and accurately collect information about health care utilization. This training also includes home study preparation and classroom learning. Again, interviewers are required to complete a certification exercise before they may conduct interviews.

Remote Interviewer Training

Once an interviewer completes in-person training, interactive remote training sessions and ongoing field management coaching are conducted throughout data collection rounds. Designed primarily for MCBS-experienced staff, remote training content focuses on the reinforcement of fundamental points in season-specific terms (e.g., data quality concerns associated with particular seasonal questionnaire sections), skill enhancement (e.g., gaining cooperation), and the refinement of fielding protocols and questionnaire administration.

Round-specific remote training includes two types of content. First, interviewers complete training modules that are designed to remind them of interview flow and critical protocols for questionnaire sections that are administered regularly. Second, interviewers complete training modules that contain in-depth content on new or revised questionnaire sections to ensure they understand the changes and administer the new or revised sections correctly.

The MCBS remote training environment is hosted virtually via an online learning management system. The remote format typically follows a series of self-guided modules of instruction, online quizzes, group calls, and mock questionnaire administration. Field managers monitor completion of remote training modules.

Field Memos and Topic Memos

The training program also includes ongoing training throughout the data collection period that closely aligns with developments in the field and challenges within the current round. Questions from the field are closely monitored and follow-up guidance is issued via weekly interviewer memos or standalone "topic memos."

6. DATA COLLECTION

MCBS data collection activities include a set of approved procedures designed to guide outreach and questionnaire administration with beneficiaries across three rounds of continuous data collection each year. In-person data collection is facilitated through a series of protocols that define eligibility for the survey, provide instruction for questionnaire administration by round and component (Community and Facility), and establish rules for how to conduct the interview within a given round. Quality control procedures are also instituted to ensure high quality data are collected.

6.1. Clearance

OMB Approval

The Office of Management and Budget (OMB) regularly approves data collection for the MCBS. For data collection activities conducted in 2015, the relevant OMB clearance was obtained on July 30, 2014 with an expiration date of July 31, 2017[33] (OMB control number 0938-0568). To address changes made to the Community questionnaire, CMS submitted two non-substantive change requests to OMB in 2015.

In June 2015, CMS submitted a non-substantive change request for OMB approval of changes to the questionnaire for Fall 2015, reflecting revisions to the Income and Assets (IAQ), Dental Utilization (DUQ), Health Functioning and Status (HFQ), and Patient Perceptions of Integrated Care (PPIC) sections. OMB clearance was received on July 7, 2015.

In September 2015, in anticipation of changes required for Winter 2016, CMS submitted a non-substantive change request for the Winter 2016 questionnaire revisions, including the addition of the Preventive Care (PVQ) section; the addition of items to the Beneficiary Knowledge and Information Needs (KNQ) section; and updated medical provider terminology in the Health Insurance (HIQ), Health Insurance Summary (HIS), KNQ, and Prescription Medicine (PMQ) sections. OMB clearance was received on December 23, 2015.

IRB Approval

The NORC Institutional Review Board (IRB) reviews and approves all MCBS data collection protocols, questionnaires, and respondent materials to ensure human subject protections were properly addressed before field data collection began. For 2015 data collection, the MCBS research protocol and consent procedures were approved by NORC's IRB in July 2014, with subsequent changes to the protocol approved through amendments and annual renewal for 2015.

---

[33] OMB most recently reviewed the MCBS in 2017 and the current clearance expires 6/30/2019.

6.2. Data Collection Process and Procedures

The MCBS data collection process includes a timeline to fulfill the continuous, three rounds of annual data collection. MCBS data collection procedures define how beneficiaries are contacted, determine when a MCBS beneficiary is eligible to participate, and include protocols designed to facilitate longitudinal data collection, establish contacting rules, and maintain beneficiary participation throughout twelve rounds over the four year period.

Data Collection Schedule and Timeline

The annual MCBS fielding schedule includes three rounds of data collection, with the Winter and Summer rounds typically lasting 16 to 17 weeks and a slightly longer Fall data collection round of 18 weeks. The fall round is scheduled as a longer data collection period to accommodate contacting and interviewing efforts with the Incoming Panel.

In 2015, the data collection schedule was altered from the typical three rounds of annual data collection to accommodate continued development and testing of the Community questionnaire required as part of the 2014 contract transition. CMS approved implementing a schedule shift through an extension of the Fall 2014 round data collection (Round 70) – instead of ending the first week of January 2015, the Fall 2014 round ended March 13, 2015. The Winter 2015 (Round 71) and Summer 2015 (Round 72) rounds were merged to create a 22-week combined round with data collection period starting March 25, 2015 and concluding August 29, 2015 (referred to as the Winter/Summer 2015 data collection round or Round 71/72). Fall 2015 data collection started August 3, 2015 and concluded January 9, 2016. Thus, data collection represented in 2015 annual files includes data collected from March 25, 2015 through January 9, 2016, with a reference period start date of January 1, 2015 for health care event data.

Sample Releases and Preloads

For a given round, MCBS data collection is structured around several case releases. This is primarily due to the cyclical nature of fielding the MCBS as a continuous longitudinal survey. For members of continuing panels, questionnaire data from the prior round need to be cleaned, edited and preloaded before a case is released into production for the next round. Continuing cases are staged and released in batches scheduled throughout the data collection round following the schedule outlined in Exhibit 6.2.1.

Exhibit 6.2.1: Typical MCBS Data Collection Round Case Release Schedule

Wave	Content	Week Number
RR.CC1	First Batch of Community Cases Released	1
RR.FAC	Single Batch of Facility Cases Released	6
RR.CC2	Second Batch of Community Cases Released	8

As part of the cyclical process of fielding and data processing, a cutover date is scheduled mid-round. It essentially batches all interviews completed prior to the cutover date in order to begin data processing and sample preparations for the first wave release in the next round.

Contacting Efforts and Outreach Rules

Given the longitudinal panel design of the MCBS, it is imperative that sampled beneficiaries engage with the study throughout the 12 rounds of data collection to minimize nonresponse bias and the impact of sample attrition over time. Recall that the MCBS data collection design no longer follows a beneficiary who misses two consecutive rounds of data collection. While beneficiaries can miss a single round, non-completion of an interview in a previous round can lead to long recall periods and less complete information collected. Various data collection strategies are used to limit respondent burden, strengthen the beneficiary's commitment to the survey and maximize response rates across rounds.

Contacting Protocols

During each case release, interviewers receive case assignments for contacting and questionnaire administration. Interviewers are trained to establish contact with respondents within a certain number and types of contact, typically starting with initial contacts to introduce the survey and gain cooperation, schedule an interview and administer the questionnaire.

Following CMS guidance, and shown in Exhibit 6.2.2, interviewers use contacting strategies that promote efficiency and ensure continuity in contacts across all beneficiaries actively fielded during a given round. The contacting effort required often corresponds to the number of rounds a beneficiary has previously participated. For example, greater effort, in terms of the number and types of contacts made, is invested in contacting the Incoming Panel beneficiaries in the first-interview Fall and second-interview Winter rounds as activities such as locating, gaining cooperation and establishing familiarity with the MCBS is often required. Contacting efforts for the 3rd through 12th interviews typically require a reduced number of attempts necessary to make contact with respondents by phone and schedule appointments to interview respondents in person.

Exhibit 6.2.2: MCBS Contacting Guidelines by Interview Round

	Interview 1	Interview 2	Interviews 3-12
Attempts made prior to initial contact with beneficiary, designated proxy, or facility staff	10 attempts	8 attempts	4 attempts
Contact attempts after initial contact to secure appointment	8 attempts	8 attempts	4 attempts
Visits to complete interview	2 visits	2 visits	2 visits

Case Management

Interviewers access their case assignments using a case management system. This system collects and displays primary contact information, contacting histories and key elements that describe case status which interviewers use to facilitate efficient outreach and questionnaire administration in a secure and standardized manner. They also use the case management system to update contact information, describe and classify outcomes of contact attempts and launch the CAPI questionnaires. This information is synchronized with central office databases for reporting and data processing tasks. See Section 4.3 for more information about the case management system.

The case management system also houses historical summaries of previously reported utilization and cost records captured during past interviews. These summaries are produced for all continuing community cases and are used by interviewers to prepare for the interview. They include information such as previously reported medicines, previously entered insurance statements, previously reported utilization without associated costs collected, and summaries of utilization events reported during the last interview.

Beneficiary Eligibility for MCBS Survey

Eligibility to participate in the survey depends upon a number of factors encountered throughout the four years of panel participation. Changes in survey eligibility are generally identified either by the interviewer while attempting to contact the beneficiary in a given round, or from Medicare program eligibility updates reported by CMS on a regular basis throughout the year. Factors that impact whether future interviews will be conducted include beneficiaries who are deceased, have lost Medicare entitlement, have relocated outside of PSU boundaries, or are no longer fielded due to *Not-in-Round* case finalization rules.

Recently Deceased. Sampled beneficiaries reported as deceased during data collection are finalized as *Complete-Deceased* at the end of the round. The standard data collection procedure for a beneficiary reported as having died at any point between the 2nd and 12th interview is to attempt an interview with a proxy in order to collect utilization and cost data between the date of the last interview and the beneficiary's date of death. A proxy completes the questionnaire in the Community setting or a final interview is completed at a facility before the case is finalized and no longer contacted in future rounds.

Starting in 2015, fielding procedures were revised to handle Incoming Panel beneficiaries reported as deceased. This change accommodates the collection of additional information as part of the "current-year enrollee" panel. The date of death reported and the beneficiary's enrollment year are key drivers for determining when an interviewer pursued a proxy interview during the first and second interviews. Following traditional MCBS fielding rules for deceased beneficiaries, any Incoming Panel beneficiary reported as deceased who became eligible for Medicare prior to the Incoming Panel year (e.g., for 2015, any Incoming Panel beneficiary who enrolled in Medicare prior to 2015) is finalized as deceased without pursuing a proxy interview. Starting in 2015, any Incoming Panel beneficiary reported as deceased who enrolled in Medicare during the same year (e.g., for 2015, any Incoming Panel beneficiary who became eligible for Medicare in 2015) is fielded for a proxy interview before being finalized as deceased. These rules apply to any Incoming Panel beneficiary who is reported as deceased at any point during the Incoming Panel year. This impacts fielding considerations in the second round Winter interview as well.

Lost Medicare Entitlement. Beneficiaries are no longer eligible for participation in MCBS after Medicare entitlement is lost. CMS provides periodic updates for beneficiaries selected to participate in the MCBS who have lost entitlement. These updates are compared with current round case management status to determine fielding procedures. If entitlement is lost while a case is being fielded as part of the Incoming Panel (first round interview), the case status is finalized as "*Ineligible for Contact."* If the beneficiary has lost entitlement during the data collection round for any continuing interview, an interview attempt is made in order to collect utilization and costs associated with the period of time when the beneficiary still maintained coverage. At the end of the continuing round, the case is finalized as "*Lost Entitlement*" and is no longer fielded in future rounds.

Beneficiaries Who Move Outside of Sampled PSUs. Consistent with fielding rules from past MCBS data collection rounds, if a beneficiary permanently moved or relocated more than 30 miles outside of MCBS sampled PSU boundaries, the case is finalized as *Moved out of Area* and not fielded in future rounds.

Case Finalization and Holdover Consideration for Fielding Next Round. Each actively fielded case is assigned a final disposition to represent the status of the case at the end of a round. Any case without a completed interview is reviewed by field management and assigned a final disposition to reflect the not-in-round status. Cases assigned not-in-round status such as final refusal or final unlocatable are no longer fielded in future rounds. The majority of beneficiaries finalized as not-in-round are no longer fielded in future rounds.

Holdover Rules for Participation. A beneficiary must have data collected for at least two-thirds of a year to be included within the annual delivery files. For data collection purposes, any respondent finalized as not-in-round for two consecutive rounds is no longer considered eligible for participation. However, to ensure participation can continue for beneficiaries unavailable in a present round but likely to participate in the future, a holdover process is used to prepare the case for fielding in the subsequent round. For example, a beneficiary could be away for an extended family visit; a beneficiary could be staying at a second home not in the area; or a beneficiary could have canceled appointments but without seeming to be a hard refusal. Cases meeting similar criteria are finalized as *"Unavailable this Round"* and are staged for fielding in the following round.

MCBS Data Collection Protocols

A primary objective of the MCBS is to collect complete information about medical care, services, and costs for each beneficiary residing in a community or a facility setting across all twelve data collection rounds. To facilitate collecting a full and complete picture of beneficiary utilization and costs, data collection protocols are used to ensure the proper mode of administration, to conduct the interview in the correct setting, and to identify rules for who respond on behalf of the beneficiary to complete the interview.

Community Questionnaire Administration

The Community questionnaire is administered in person. Longstanding MCBS protocols have required that Incoming Panel beneficiaries be contacted and interviewed in person. This approach ensures survey legitimacy is established early on and allows the interviewer to establish rapport with the respondent, provide context for future rounds, and introduce materials in support of future rounds.

A key goal of continuing interviews involves associating health care events with costs and payments. In preparation for the future rounds, interviewers provide respondents with a calendar and instructional aid that reminds them to document medical events and save any Medicare or insurance statements and any other health care-related paperwork received after the date of the current interview. During the subsequent round, interviewers review calendars with respondents, as well as sort and match any hardcopy documentation associated with past reported medical events such as Medical Summary Notices (MSNs), explanation of benefits (EOBs) and other supplemental insurance forms, and medicine summaries. Interviewers are trained to match these documents into charge bundles to ensure streamlined entry within the questionnaire (see Section 4.1 for more information on how these statements are used during the Cost Series).

The only MCBS questionnaire that is administered by telephone administration is the shorter 12th round interview for Community respondents exiting the survey. This final interview does not collect utilization and cost information, so in-person documentation matching and sorting is not required.

Facility Component Interviewing

If a beneficiary spent time in both the community and a long-term care facility during a given round of data collection, both community and facility interviews may be administered to ensure that continuous records are obtained for the entire reference period. Prior to conducting a facility interview, a potential facility must be screened to ensure the facility meets the MCBS facility definition.

MCBS Definition of a Facility

For the MCBS, a facility interview is conducted when the beneficiary resided in a long-term care or other residential facility that meets the following conditions.

The facility must be a place or unit of a larger place with <u>three or more</u> beds, and meets one of the following:

- Is certified by Medicare as a Skilled Nursing Facility (SNF); or

- Certified by Medicaid as a Nursing Facility or an Intermediate Care Facility for the Mentally Challenged; or
- Is licensed as a Personal Care Home, Board and Care Home, Assisted Living Facility, Domiciliary Care Home or Rest Home by a state or local government agency; or
- Provides 24 hours a day, 7 days a week supervision by a person willing and able to provide personal care; or
- Provides personal care services to residents (personal care may include assistance with eating, dressing, walking, preparing meals, etc.).

If a facility does not meet the above definition, or if the beneficiary does not reside in the section of the facility that provides long term care, then a Community questionnaire is instead administered to collect the data.

Most beneficiaries who reside in a place that meets the MCBS definition of a facility live in a type of nursing home. Other qualifying facilities include institutions for those with mental disabilities, domiciliary or personal care homes, retirement homes, mental health facilities, assisted living, board and care homes, rehabilitation facilities, and group homes.

Institutions such as jails and prisons do not meet the MCBS facility definition. The Facility Screener and the Facility Questionnaire (FQ) section, the first section within the Facility Instrument, are used to confirm that a facility meets the MCBS definition. The Screener and FQ work in tandem to determine whether a case is eligible for the Facility component.

Facility Screener

When an interviewer learns that a beneficiary who was previously residing in the community has moved into a facility, or a beneficiary who was residing at a facility has moved to a new facility, the interviewer determines whether the new facility meets the MCBS definition of a facility and therefore is eligible for the Facility interview.

As a first step in determining eligibility for the facility interview, the interviewer administers a Facility screener over the phone to a facility contact. The Facility Screener serves to confirm the beneficiary has lived in the facility during the reference period, identifies the current location of the beneficiary, and verifies the location of the facility and relevant contact information.

Facility Instrument Administration

Unlike in the Community component, interviewers never directly administer the questionnaire to the beneficiary during a Facility interview. Instead, the interviewer administers the questionnaire to staff at the facility, referred to as “facility respondents”, who answer questions about the beneficiary. It is common for field interviewers to interview more than one person at the facility because different staff at the facility have the most complete information for specific sections of the questionnaire.

Much of the content of the Facility interview can be found in medical documentation. Therefore, facility staff may refer to records, such as the beneficiary's medical chart, during the interview. Further, facility staff may allow the interviewer to abstract responses directly from medical records. The extent of abstraction conducted varies greatly by instrument section, facility structure, and number of events occurring at the facility on the day of the interview. Exhibit 6.2.3 shows the percentage of cases for which interviewers reported abstracting data by section as reported at the conclusion of each questionnaire section for which such data are available.[34]

Exhibit 6.2.3: Interviewer-Reported Abstraction Rates by Facility Instrument Section

Questionnaire Section	Reported Abstraction in Winter/Summer 2015	Reported Abstraction in Fall 2015
Background (BQ)	13%	12%
Prescribed Medicine (PM)	52%	54%
Residence History (RH)	9%	8%
Health Status (HS)	26%	23%
Use of Health Services (US)	20%	23%

Crossover Definitions and Procedures

If a beneficiary spends time in both the community and a long-term care facility during a given round of data collection, both Community and Facility interviews are staged for administration to ensure that continuous records are obtained for the entire reference period. Crossovers are cases that have moved into a new setting since the last interview.[35] In a crossover situation, because the beneficiary has spent part of the reference period in more than one setting, interviewers complete two separate questionnaires to collect data from both locations.

Incoming Panel cases in Winter/Summer 2015 had a different policy that depended on when the beneficiary entered the new component and when s/he gained Medicare entitlement. All other crossover cases in their 3rd-11th interviews follow the crossover procedures outlined below.

Community-to-Facility Crossover. When a contact attempt with a continuing Community beneficiary leads to the discovery that the beneficiary moved into a facility since the last interview, a Community-to-Facility crossover occurs. An interviewer first attempts to administer the Community interview to a proxy followed by administering the facility screener to staff at the facility where the beneficiary is residing. Once the facility screener confirms that the facility meets the MCBS definition, an appointment is scheduled to conduct the facility interview. In Winter/Summer

---

[34] Data regarding abstraction in the Expenditures (EX), Insurance (IN), and Facility Questionnaire (FQ) sections are not collected.

[35] Crossovers do not include respondents that have moved, but remained within the Community setting.

2015, a systematic automated crossover process for staging a Facility questionnaire was implemented allowing both the Community and Facility questionnaires to be fielded within the same round.

Facility-to-Community Crossover. When contact with a facility where a continuing beneficiary was residing during the last interview indicates that the beneficiary moved back to the community setting, a Facility-to-Community crossover occurs. An interviewer administers the Facility interview and obtains information such as the date the beneficiary left the facility as well as the beneficiary's current community residence.

Facility-to-Facility Crossover. When contact with a facility where a continuing beneficiary was residing during the last interview indicates that the beneficiary moved to another facility since the date of the last interview, a Facility-to-Facility crossover exists. An interviewer first administers the Facility interview with the original facility to cover utilization and costs from the date of the last interview through the time of the move into the new facility. The interviewer then collects the required facility screener information for the case to be fielded in the second Facility setting. This second facility will be contacted in the following round.

Proxy Interviews and Assistants

Beneficiaries often require assistance in providing the detailed information needed to accurately respond to survey items. During the course of data collection, the beneficiary may designate a proxy to participate in the interview on his or her behalf or an assistant to provide help when responding to specific survey questions.

Proxies and Assistants

A proxy is a person, generally designated by the beneficiary, who is sufficiently familiar with the beneficiary's health care events and costs and responds on behalf of the beneficiary. In addition, a proxy completes a Community interview when a beneficiary is no longer able to participate, including when a beneficiary died since the date of the last interview, or has entered a Facility setting. In 2015, the percent of interviews completed by a proxy ranged from 11-13% depending on the round.

An assistant helps the beneficiary answer specific questions, but unlike a proxy, an assistant does not answer all questions on behalf of the beneficiary. The assistant is chosen by the beneficiary to help in situations where the beneficiary could respond to the interview as long as he/she received some help from another knowledgeable person, such as instances where a spouse or partner manages the Medicare statements for the household or maintains a calendar of medical visits and appointments. The percent of interviews completed with the help of an assistant in 2015 ranged from 16-21%, again depending on the round.

Criteria for Proxy Selection

During Community Questionnaire administration, all beneficiaries are asked to identify a person or persons best able to provide information about health care visits and the costs of any health care the beneficiary may receive should the beneficiary not be able to complete a future interview. For Continuing round interviews, the named proxy is in the case management system, along with information indicating if a proxy completed the interview in the prior round. Community interviews conducted with proxies follow a slightly different path than those administered directly to the beneficiary (see Section 4.1 for the Community Questionnaire flow).

When initial contacts with Incoming Panel beneficiaries suggest possible comprehension or physical impairments that would make the interview difficult, interviewers work with their managers to determine if an assistant or proxy is necessary, and who an appropriate person would be to serve as a proxy or assistant.

Interviewing Languages

The Community questionnaire is programmed for administration in English or Spanish. The Facility Instrument is available for administration in English. Approximately 5% of Community interviews were conducted in Spanish in 2015. This includes interviews conducted in Puerto Rico, all of which are completed in Spanish.

Bilingual FIs are trained to administer the Community questionnaire in both English and Spanish. The language of administration is captured within the questionnaire. In rare instances in which the beneficiary speaks a language other than English or Spanish, the interview is conducted in English with an English-speaking proxy or assistant acting as an interpreter for the beneficiary.

Questionnaire Breakoffs

Interviewers are able to suspend the interview prior to completion while administering both the Community and Facility questionnaires. This break-off feature provides flexibility to address schedule constraints, technical issues, and other extenuating circumstances that prevent completion of the interview in one sitting. Once restarted, the CAPI resumes at the screen of the last question administered. If a questionnaire is broken off, it must be fully administered before the end of the round to count as a completed interview. If the suspended questionnaire is never completed, it is finalized as a *Final Breakoff* at the end of the round.

6.3 Data Collection Results

An interview is considered to be complete once administration of all questionnaire sections to the respondent has concluded, the Interviewer Remarks Questionnaire (IRQ) is completed and data are fully transmitted. For 2015, the length of community interviews in Winter/Summer 2015 was slightly longer than expected – a mean of 85 minutes – in part due to the longer reference period of the combined round. The length of interview for Fall 2015 was also longer than expected – about 90 minutes for the Incoming Panel and approximately 120 minutes for the Continuing beneficiaries.

The longer than usual length of the Fall 2015 interview was due to the addition of the Income and Assets section as well as the expansion of the Usual Source of Care to include Patient Perceptions of Integrated Care.

Exhibit 6.3.1 provides the count of completed interviews by round and component for 2015. Detailed information on response rates can be found in Chapter 9.

Exhibit 6.3.1: 2015 Completed Interviews by Component

Round	Component	Completed Interviews
Winter/Summer 2015	Community Only	12,172
	Facility Only	807
	Both	29
	Total	**13,008**
Fall 2015	Community Only	12,437
	Facility Only	1,001
	Both	22
	Total	**13,460**

6.4 Data Collection and Quality Control

To ensure the collection of high quality data, several quality control procedures are conducted including systematic review of questionnaire data and case management paradata, follow-up contacts with respondents, and ongoing interviewer coaching. Systematic review of interview recordings and ride-along observations of in-person interviews are used to directly observe interviewer interaction with beneficiaries and provide feedback. Verification phone calls and review of survey data are also conducted to validate interviewer performance.

The systematic monitoring and evaluation of interview performance and verification is primarily conducted via digital computer-assisted recorded interview (CARI) recordings. A subset of questionnaire items are recorded with respondent consent. By listening to a random sample of CARI recordings, supervisors identify areas where interviewers require correction in administration, stress the improvement of interviewer techniques to add clarity or minimize potential bias, and emphasize standardization in approach and administration. Any serious deviations from protocol or data quality concerns are reviewed for corrective action in consultation with field management.

Data review procedures are also enacted to identify any systematic CAPI issues resulting from the data collection effort. In 2015, data review procedures consisted of two components: review of survey data within the preload data cleaning process, and review of metadata to assess interviewer performance. Because the continuing interview by design is highly dependent upon data collected in prior rounds, a multistep cross-team process is used to review questionnaire data prior to preloading for the next data collection round (see Section 7.1). The data cleaning process informs future questionnaire development as well as additional training and follow-up.

Finally, field managers periodically contact respondents throughout the round to verify the interview was conducted, confirm the interviewer was present, and collect administration information. When necessary, field managers use CARI reports and data review feedback that indicate potential quality issues to prioritize follow-up contacts in order to collect additional information for coaching purposes.

7. DATA PROCESSING AND DATA DELIVERY

Longitudinal data collection requires both interim and final post-processing of the data in order to prepare them for release. These activities include data editing, for both preloading subsequent round instruments and final file production, data concatenation and reconciliation for the annual data products including the Survey file and Cost Supplement files, and the development of other post-processing inputs to the files. This chapter describes both the data editing process and the annual data concatenation and reconciliation process.

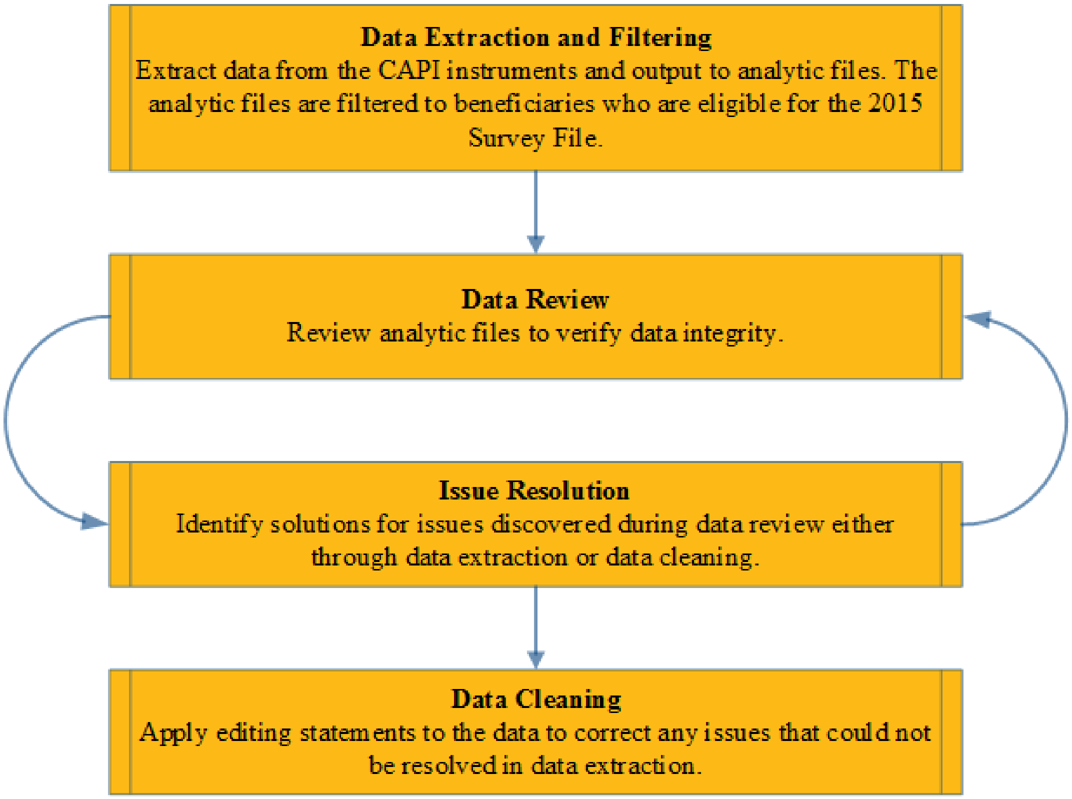
7.1 Data Editing Overview

CAPI data are reviewed and edited for three primary purposes: Community and Facility Questionnaire preloads, the 2015 Survey File (formerly Access to Care or ATC), and the 2015 Cost Supplement File (formerly Cost and Use or CAU). The same types of data review and editing protocols are used for each effort, with different source instruments and editing protocols. This section provides an overview of data review and editing procedures. The sections that follow will provide further description of the individual editing efforts.

Process Description

Exhibit 7.1.1 illustrates steps in and iterative nature of the data review and editing process.

Exhibit 7.1.1: Data Review and Editing Process



Data Extraction and Filtering
Extract data from the CAPI instruments and output to analytic files. The analytic files are filtered to beneficiaries who are eligible for the 2015 Survey File.

Data Review
Review analytic files to verify data integrity.

Issue Resolution
Identify solutions for issues discovered during data review either through data extraction or data cleaning.

Data Cleaning
Apply editing statements to the data to correct any issues that could not be resolved in data extraction.

Data Extraction and Filtering. At the conclusion of data collection in each round (Fall 2014, Winter/Summer 2015, and Fall 2015) data are extracted from the raw Community and Facility CAPI questionnaires and transformed into SAS analytic files for further processing. This extraction includes the development of appropriate questionnaire metadata. The definition of the analytic tables used for processing reflects the data structure required to transport data preloads from round to round and the preparation of inputs for the final data files. Prior to data review, the individual records and associated analytic files are limited to beneficiaries who are deemed eligible for the appropriate data collection or data product.

Data Review and Issue Resolution. Given the complexity of the data structure, the analytic files undergo column and row checks to confirm each individual analytic file is structurally sound. Column checks include confirming that all necessary variables are on the file; checking variable attributes, and identifying high rates of missingness or out of range values. Row checks include confirming the inclusion of expected beneficiary IDs and checking for duplicate or missing linking variables. Structural issues discovered during this process may reinitiate the data extraction process or may be resolved in data cleaning.

Logic and reasonableness checks follow for each analytic file. Logic checks are used to verify that the questionnaire worked as expected, particularly with respect to questionnaire routing and skip logic. The complexity of the event and cost questionnaire sections require particular attention to the

CAPI routing routines that generate analytic tables that are specific to these portions of the questionnaire. Errors identified during logic checking result in two types of data edits: flagging values that were incorrectly skipped or setting incorrectly populated values to null to indicate a valid missing.

Furthermore, unreasonable or impossible values are identified and checks for values that are not explicitly disallowed by the questionnaire are identified for reasonableness. For example, in the Community Questionnaire, male beneficiaries should not report female-only conditions, like cervical cancer. Based on the results of this data review, edits are developed to correct the errors during data cleaning.

Data Cleaning. Once the data review and issue resolution steps are complete for each analytic file, data cleaning routines are implemented. During data cleaning, edits are applied to the analytic file and additional quality control (QC) is conducted to ensure that the edits are applied correctly.

7.2 Preload Editing and File Production

This section describes Community and Facility Questionnaire preload production, including the purpose of preloads, examples of preloaded variables, and a general description of timeline and processes. The preload process uses questionnaire data from previous rounds' interviews and populates the Community and Facility CAPI Instruments to help drive data collection in the subsequent round. Preloaded data serve to both forestall asking MCBS panel members the same questions in subsequent rounds and to act as the basis for collecting additional information about a medical event, insurer, or associated financial cost or payment. As the data must be loaded into an active CAPI instrument available to interviewers, it requires that the preload data are in a form that is recognized by the case management system, which supplies it to the Community and Facility Questionnaires in the field. Preloaded information is used to determine questionnaire routing and text fills.

For example, if a beneficiary previously reported having ever smoked a cigarettes in his/her lifetime, the questionnaire can then use this information in a subsequent round to probe if the respondent is still smoking. The logic within the questionnaire that determined whether such a question was asked in the next round is driven by preload variables set during the preload process. Examples of preloaded data included information on health plans, medical events, insurance claims, prescription medicines, household members, facility characteristics, and facility stay history.

Preloads generally fall into two categories: direct response data and derived variables. Direct response data are raw questionnaire responses generated in one round that are passed through to the next round. For example, the list of a beneficiary's medical care providers are passed from one round's Community Questionnaire to the next via the preload process. Similarly, facility name and address are passed from one round's Facility Questionnaire to the next.

Derived variables require modification of the source data before being preloaded into the next round. Some modifications are quite complex and many derived variables have a significant impact on questionnaire functioning. Examples of derived variables include sample type assignments, Facility and Community Questionnaire reference dates, and the reason a cost is sent through Charge Payment Summary (CPS reason) (see Section 4.1 for more details on this questionnaire section).

Community and Facility Preload Timeline

Preload creation is an ongoing process with a rigid schedule as each round lasts only four months. The MCBS is continuously fielded, thus, the preloads must be ready for the next round of data collection to be conducted.

Each round of data collection in 2015 contained four waves of sample release that must include Community preloads. The Continuing Community Release 1 beneficiaries completed the prior round's Community Questionnaire within the first half of the data collection period while Continuing Community Release 2 beneficiaries completed in the second half. There is one Continuing Facility Release each round, which includes beneficiaries that either completed the prior round's Facility Questionnaire or crossed over from the Community Questionnaire to the Facility Questionnaire because the beneficiary continues to live in a facility or has entered a facility in the new round. Timing for these wave releases is described in Section 6.2 (Exhibit 6.2.1).

Community and Facility Preload Process Description

The Community and Facility Questionnaire preload creation processes consist of five steps: data extraction and filtering, data review, issue resolution, data cleaning, and rollover. The first four steps were described previously. The eligible population for each subsequent round is determined by examining case dispositions in the current rounds and items extracted for review and editing reflect the round specific preload requirements.

The final phase of preload creation, which was not described in the introduction, is the rollover process. In the preload process, after data review and editing occur, datasets are constructed with data required for preloading. Key items set during the rollover process are the derived variables that assign sample type, reference dates, and CPS reason. Sample type assignment is based on previous interview history, including whether respondents missed the previous interview, were ever in a facility or the community, or are in their first year of the MCBS. This information is used to determine which questionnaire sections and items are administered and set the reference dates for questionnaire items. Reference dates are used in the Community and Facility Instruments to define the time periods about which data will be collected in the upcoming round. There are a number of reference dates that are derived from the dates of the respondents' prior interviews. CPS reason determined which medical costs are collected in the Community Questionnaire based on whether the respondent has a billing statement for that item and whether the total charges were accounted for in previous rounds. The rollover process, which is designed to ensure that all of the preload data is loaded properly into the questionnaire, occurs before every sample load in a round and between rounds.

Thorough quality control steps, including ensuring the data types, dates, and variable definitions are appropriate, are conducted to ensure that preloaded data are successfully created per the round based specifications. These steps are designed to check on the transformation of data from SAS files into case management database tables, which are necessary to make the preload data available in the instrument used by FIs in the field. This QC is critical, as the preload data needs to be in the specified format acceptable to the case management system, which then makes the preload data available to be called in the Community and Facility Questionnaires for the upcoming round.

7.3 MCBS 2015 Survey Data File

The 2015 Survey File release is built from 39 analytic data files encompassing Community and Facility data collection from four rounds of data including Winter/Summer 2015, and Fall 2015, Winter 2016 and Summer 2016. These files are input into CMS processes that generate the final data files available to the public. Further detailed descriptions of these final Survey Data files are available in the 2015 Data User Guide. This section describes the eligibility criteria for analytic file, file preparation, and file contents.

File Eligibility Criteria. Each analytic file that provides input to the final survey data file release contains data collected in Winter/Summer 2015 and Fall 2015 for all completed, eligible Community or Facility interviews. The criteria for inclusion in these analytic files include: beneficiaries continuously residing in the Community or Facility, beneficiaries who move between a Facility and the Community, data which are collected from proxy respondents for deceased beneficiaries, or data from individuals who lost entitlement to Medicare. A beneficiary only needs to have completed a Community or Facility interview in one of the data collection rounds of interest to be included in these analytic files.

File Contents

Community

This includes two subcategories of the Community analytic files. The Community Core Questionnaire analytic files contain data collected in questionnaire sections critical to the purpose of the MCBS. Core data are collected in each round of an annual data collection cycle. The Community Topical Questionnaire analytic files contain data collected in questionnaire sections that cover special interest issues. Topical data are collected only in selected rounds. See Chapter 4 (Exhibits 4.1.4 and 4.1.6) for a list of the Community Questionnaire sections included in each data file.

Community Core Questionnaire

The Community core content analytic files included in the 2015 Survey File contain information about access to medical care, health status and functioning, health insurance plans, medical providers, and income and assets. The Survey File analytic files do not include survey-reported cost, healthcare utilization, or case management data.

Community Topical Questionnaires

The Community Questionnaire include sections that are focused on specific topics of interest, such as mobility of beneficiaries and preventive care and drug coverage. The 2015 Survey File contains data from some topical sections that were administered in the Winter 2016 (Round 74) and Summer 2016 (Round 75) round but have reference periods for 2015. These files are processed in combination with the 2015 Survey File deliveries and as a result, topical analytic files were considered part of the MCBS 2015 Survey File.

Facility

The Facility analytic files contain information about access to medical care, health status and functioning, health insurance plans, facility characteristics, and beneficiary characteristics. The Facility analytic files do not include cost, healthcare utilization, or case management data.

7.4 MCBS 2015 Cost Supplement File

The Cost Supplement File data include information on beneficiaries' medical events occurring in 2015, and the cost of those events. The MCBS Cost Supplement File is described in the 2015 Data User's Guide.

Substantial post-processing occurs to the questionnaire items that create health care events, the costs and payments associated with those events, and the source of payments. This is the result of the way in which the annual data collection occurs. Five processes are used to create the inputs to the final data files. The five processes build annualized files, define eligibility for the Cost Supplement File, establish insurance coverage across the year and create events that are linked to defined payers and the cost of the services provided. The first four analytic processes are inputs to the claims match process that return matched events for additional post-processing and imputation. The final process, the facility stay file, combines all the steps already described for the community questionnaires and ads the claims match into a single step. The facility stay process then generates data files for the production of the Cost Supplement File release.

These five processes (reference period, insurance timeline, event cost consolidation, prescription medicine file, and facility stay file) are described below. The shared goal of all of these interim analytic steps are to combine data across rounds, annualize eligibility for data release, and create analytic products that can be consumed in the context of the final file production. These interim data products are not part of the final Survey File or Cost Supplement File releases.

Eligible Population

Each of the Cost Supplement File data analytic products described have a different eligible population of beneficiaries that depend on the goal of the product.

Reference Period is run for all beneficiaries that may have had an interview in 2015, and it is used to identify the population of beneficiaries eligible to be included in the Event Cost Consolidation and

Prescription Medicine File. The population identified by Reference Period represents a subset of the population assigned an Ever-Enrolled weight.

Insurance Timeline in 2015 was produced for the same population as was assigned an Ever-Enrolled weight.

The Facility Stay File population in 2015 included any beneficiaries completing one or more Facility Questionnaire interviews covering residence in an MCBS-eligible facility for one or more days in 2015.

Reference Period

Reference Period uses case management data to define time periods in 2015 covered by Community and Facility survey data. Reference Period 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 had complete 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 Event Cost Consolidation and Prescription Medicine File data products.

Insurance Timeline

Insurance Timeline creates a calendar history of a beneficiary's insurance plans and types of insurance coverage. The process pulls together health insurance plan data from the Community Questionnaire, Facility Instrument, and administrative records. Insurance plan timelines are constructed independently across these three data sources. Plans that are identical across data collection periods are collapsed into one record, with each time period identified as having definite or possible coverage by the plan. Plans identified as "Medicare HMO" in the Community Questionnaire data are linked to Medicare Advantage (MA) plans in the administrative and claims data. Finally, the three timelines from each data source are concatenated. The resulting dataset allows these timelines to be examined independently or together to understand insurance coverage in the calendar year for each beneficiary. Plan coverage data from the Insurance Timeline are used downstream to define potential sources of payment in the Event Cost Consolidation process as well as used to construct monthly insurance coverage records for each beneficiary.

Event Cost Consolidation

Event Cost Consolidation creates a file containing health care events and their associated costs, payments, provider information, and dates of service for all health care utilization reported by or on behalf of beneficiaries residing in the community. The process matches events to reported periods of insurance coverage as summarized by Insurance Timeline to identify possible and definite sources of payment for each event. Reported charges and payments are matched before being appended to the file of events. The process then applies global editing rules to resolve partial charges and charges with incomplete cost information. Finally, records for recurring events are replicated to represent repeated instances of these events. The resulting dataset of consolidated event and cost information are used to match survey-reported events to Medicare claims. These matched results are the inputs to the Prescription Medicine and non-Prescription Medicine Imputation processes and the final Cost Supplement Files.

Prescription Medicine File

The Prescription Medicine (PMED) file is a list of all prescription medicines that are collected by the MCBS. For 2015, the list included every combination of prescription medicine names, forms, and strengths provided by MCBS respondents to the Community interview during interviews conducted in 2015 (including the Fall 2014 round through the Winter 2015 round). It includes both medicines that were reported by MCBS respondents for the first time during one of the interviews conducted in 2015 and refilled medications that were originally reported earlier, but updated as being currently prescribed during 2015. It only includes medicines that were reported during the Community Questionnaire administration for beneficiaries who were eligible to be included in the Cost Supplement File.

A number of new and enhanced cleaning steps were developed for the first time in the 2015 PMED file. These cleaning steps included fixing common misspellings and other common errors in the verbatim fields entered by interviewers and standardizing the spacing, punctuation, and other formatting of the prescription medicine information. The purpose of the cleaning was to simplify the subsequent CMS process of matching the PMED list to the First Databank list of prescription medicines, and eventually to administrative claims information.

The process of creating the PMED file includes assembling a full list of all beneficiaries' reported prescribed medicines for 2015 from the Community Questionnaire, de-duplicating it, developing and implementing cleaning rules, and then de-duplicating the list again after the cleaning process was complete. For the 2015 data, the original list of prescription medicines included over 60,000 records. After cleaning and de-duplicating again, the final file included just over 45,000 unique medicines.

Facility Stay File

The Facility Stay File summarizes data related to facility characteristics, costs and payments, and health care utilization for interviews conducted on behalf of beneficiaries residing in facilities. The process brings in data from the Facility Questionnaire and reconfigures the data to create one record per facility stay during the calendar year. Medicare Claims data for inpatient hospital visits and skilled nursing facility visits are matched to Facility Instrument data to provide more accurate reporting of Medicare payments. Three imputation routines are applied within the context of the Facility stay process to remedy missing data issues with payments as well as edit outliers and other anomalies.

8. WEIGHTING AND IMPUTATION

8.1 Overview

Weighting and imputation are used in surveys to enhance the usability of the data for analysis and increase the accuracy of resulting estimates. Weights are calculated to reduce potential nonresponse and sample coverage bias, ensuring that the sample is representative of the population of interest. They are especially important when particular sampling methods are in place, such as stratification, cluster sampling, and oversampling of particular populations. The MCBS employs all of these sampling methods; weights then account for the resulting differences in probabilities of selection as well as nonresponse, and also calibrate to control totals using post-stratification. Imputation is used to replace missing values of survey variables with admissible complete values and create data where they were not actually collected, allowing for the retention of observations for statistical analysis that would otherwise be excluded. MCBS imputation falls under two umbrellas that focus on imputing monetary amounts: Income and Asset (IA) imputation, and Event, Payer, and Cost imputation, which includes imputation for Prescription Medicine (PM) and Non Prescription Medicine (Non PM) events and costs. The weighting and imputation methods used for the MCBS are described in detail below.

8.2 MCBS Weighting Procedures

Overview

Weighting activities for the 2015 data year consist primarily of four main stages. The first is the initial weighting stage in which the members of the Incoming Panel are given base weights, and these weights are then raked to population control totals and adjusted for nonresponse at the first interview (Fall 2015). The remaining three stages of weighting each lead to delivered weights files. These are the Survey File weights, the Cost Supplement weights, and the weights for topical questionnaire sections. A listing of all of the weights for the MCBS is presented in Exhibit 8.3.1.

Process

Initial weighting requires receipt of the final combined enrollment data extracts and the finalization of the interview dispositions in the fall round of the data year (i.e., Round 73 for the 2015 data year). Survey File weighting follows initial weighting. Cost Supplement File weighting requires completion of the Survey File weighting process and the Reference Period process. Topical questionnaire modules 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.

2015 Initial Weighting

In the initial weighting stage, the initial nonresponse adjusted weights for the incoming panel of Medicare beneficiaries, which for the 2015 data year is referred to as the “2015 Panel” or the

"Incoming Panel" are derived. First, base weights are calculated based on the probabilities of selection for the beneficiaries in the panel and 100 replicate weights for use in variance estimation are created. Then, these weights are raked to population control totals. Finally, the weights are adjusted for nonresponse at the first interview in Fall 2015.

Exhibit 8.2.1: 2015 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	CS1YRWGT	CS1YR001-CS1YR100	Continuously-enrolled from 1/1/2015 through the fall of 2015
Survey File	Ever Enrolled Cross-Sectional Weights	EVRWGTS	EEYRSWGT	EEYRS001-EEYRS100	Ever enrolled for at least one day at any time during 2015
Survey File	Continuously Enrolled One-Year Longitudinal Weights	Will not be released in 2015	-	-	Continuously-enrolled from 1/1/[2015-1] through the fall of 2015
Survey File	Continuously Enrolled Two-Year Longitudinal Weights	LNG3WGTS	L3YRSWGT	L3YRS001-L3YRS100	Continuously-enrolled from 1/1/[2015-2] through the fall of 2015
Survey File	Continuously Enrolled Three-Year Longitudinal Weights	LNG4WGTS	L4YRSWGT	L4YRS001-L4YRS100	Continuously-enrolled from 1/1/[2015-3] through the fall of 2015
Cost Supplement File	Ever Enrolled Cross-Sectional Weights	CSEVRWGT	CSEVRWGT	CSEVR001-CSEVR100	Ever enrolled for at least one day at any time during 2015
Cost Supplement File	Two-Year Longitudinal Weights	Will not be released in 2015	-	-	Enrolled on or before 1/1/[2015-2] and still enrolled at any time during 2015
Cost Supplement File	Three-Year Longitudinal Weights	Will not be released in 2015	-	-	Enrolled on or before 1/1/[2015-3] and still enrolled at any time during 2015
Survey File Topical Module	KNQ Continuously Enrolled	MCREPLNQ	KNCWT	KNC1-KNC100	Continuously enrolled in 2015 and still alive, entitled, and non-institutionalized in Winter 2016
Survey File Topical Module	KNQ Ever Enrolled	MCREPLNQ	KNEWT	KNE1-KNE100	Ever enrolled in 2015 and still alive, entitled, and non-institutionalized in Winter 2016

Limited Data Set	Description	Segment	Full-Sample Weight	Replicate Weights	Population
Survey File Topical Module	IAQ Continuously Enrolled	INCASSET FOODINS	IACWT	IAC1-IAC100	Continuously enrolled in 2015 and still alive, entitled, and non-institutionalized in Summer 2016
Survey File Topical Module	IAQ Ever Enrolled	INCASSET FOODINS	IAEWT	IAE1-IAE100	Ever enrolled in 2015 and still alive, entitled, and non-institutionalized in Summer 2016
Survey File Topical Module	PAQ Continuously Enrolled	PNTACT	PACWT	PAC1-PAC100	Continuously enrolled in 2015 and still alive, entitled, and non-institutionalized in Summer 2016
Survey File Topical Module	PAQ Ever Enrolled	PNTACT	PAEWT	PAE1-PAE100	Ever enrolled in 2015 and still alive, entitled, and non-institutionalized in Summer 2016
Survey File Topical Module	RXQ Continuously Enrolled	RXPARTD	RXCWT	RXC1-RXC100	Continuously enrolled in 2015 and still alive, entitled, and non-institutionalized in Summer 2016
Survey File Topical Module	RXQ Ever Enrolled	RXPARTD	RXEWT	RXE1-RXE100	Ever enrolled in 2015 and still alive, entitled, and non-institutionalized in Summer 2016

Full-sample and Replicate Raked Base Weights

A full-sample base weight is derived for all beneficiaries in the 2015 Panel. The base weight is equal to the inverse of the beneficiary's overall probability of selection and reflects probabilities at the PSU, SSU, and beneficiary (USU) sampling stages. Let $\pi_{k|ij}$ be the conditional probability of selection for beneficiary k given the PSU i and the SSU j, such that $\pi_{k|ij} = \rho_{1ak|ij}$ for beneficiaries in the Hispanic sampling stratum and age group a, and similarly equals $\rho_{-1ak|ij}$ and $\rho_{2ak|ij}$ for beneficiaries in the non-Hispanic and Puerto Rico sampling strata, respectively, as described in Chapter 3. Then, for all selected beneficiaries, the base weights are defined by

$$W_{1ijk} = \frac{20}{\pi_i \pi_{j|i} \pi_{k|ij}}$$

where π_i is the probability of selection for the PSU, $\pi_{j|i}$ is the conditional probability of selection for the j-th SSU given the PSU, and $\pi_{k|ij}$ is the conditional probability of selection for the k-th beneficiary in the 5-percent enrollment data extract given the PSU and SSU.

Then, one hundred replicate base weights are derived from the full sample base weights, using the variance stratum, and the variance unit of the beneficiary. The variance strata and variance units are derived from the PSUs and SSUs used for sampling. For sampled beneficiary ijk as described above, the α = 1, …, 100 replicate weights for BRR estimation are defined by

$$W_{1ijk\alpha} = \begin{cases} \{\tau(H_{h\alpha} + 1) + (1 - \tau)(1 - H_{h\alpha})\}\, W_{1ijk} & \text{if in stratum } h \text{ and unit 1} \\ \{\tau(1 - H_{h\alpha}) + (1 - \tau)(H_{h\alpha} + 1)\}\, W_{1ijk} & \text{if in stratum } h \text{ and unit 2} \end{cases}$$

where H_{ha} is the associated element in a 100x100 Hadamard matrix. For calculation purposes, this can be written as

$$W_{1ijk\alpha} = 2\left[\tau\delta_{j\alpha} + (1 - \tau)(1 - \delta_{j\alpha})\right]W_{1ijk}$$

where τ is a compositing factor between zero and one, δ_{ja} is a 0-1 indicator of whether the beneficiary is in replicate half-sample a as determined by the value of H_{ha} , and W_{1ijk} is the base sampling weight for the beneficiary. A value of $\tau = 0.85$ is used, continuing the practice used in prior MCBS years.

The full-sample and replicate base weights are then adjusted in such a way that the sum of the weights for various demographic domains are equal to pre-determined control totals based on the enrollment data extracts, through a process called "raking." The final enrollment data 5-percent extract, received in January 2016, contained additional records for beneficiaries that became eligible near the end of 2015. Due to the timing of this file, these newly-added beneficiaries were not subjected to sampling and could not be included in the 2015 Panel. This small amount of effective population undercoverage is adjusted for in this raking step. Thus, even though those beneficiaries are not eligible for sampling, they are counted in the population totals. This ensures that the weights for the 2015 Panel sum to the correct population total.

The raked full-sample weight is defined by

$$W_{2ijk} = \varphi_{ijk} W_{1ijk}$$

where φ_{ijk} is the raking step adjustment factor for beneficiary ijk. The raking process calibrates the weights by adjusting them to match the control totals for the first raking dimension, then for the second raking dimension, then for the third dimension, and so on, iterating until the weights perfectly match the control totals in all dimensions. The four dimensions used at this raking step are

3. Age Group (5-level) × Sex (2-level) × Race (2-level)
4. Census Region (4-level) × Age Group (5-level)
5. Metropolitan Status (2-level) × Age Group (5-level)
6. Accretion year (6-level; year of enrollment in Medicare)

This adjustment, and all adjustments mentioned in the remainder of this chapter, are made both to the full-sample weights and the 100 replicate weights.

Initial Nonresponse Adjustments

The raked base weights for the 2015 Panel are then adjusted for nonresponse at the first interview in Fall 2015. The response statuses in Fall 2015 are determined, where a respondent is a beneficiary that is alive and entitled and completed the Fall 2015 interview. Nonresponse adjustment cells are constructed prior to performing the adjustment. First, the beneficiaries are divided into four primary adjustment cells: alive community, deceased community, facility, and Puerto Rico.

Separately within each of these main adjustment cells, response propensity models are fit using logistic regression to model the probability of response at Fall 2015 as a function of covariates derived from multiple sources. These include county-level American Community Survey (ACS) estimates, tract-level ACS estimates, county-level physician fee schedules, rural-urban and Metropolitan Statistical Area (MSA) information, and administrative and claims data at the beneficiary level. Generally, the covariates are selected into the logistic regression model using stepwise selection procedure with an entry p-value of .10 and a stay p-value of .15. Using the predicted response probabilities, beneficiaries are grouped into cells of approximately 100 each. Separately within each of these cells, a ratio adjustment is performed to distribute the weights of the nonrespondents to the respondents, where the adjusted weights are defined by

$$W_{3ijk} = \left(\frac{\sum_{ijk} W_{2ijk}}{\sum_{ijk} I(ijk \in R) W_{2ijk}} \right) W_{2ijk}$$

where $I(ijk \in R)$ is a 0-1 indicator function indicating whether beneficiary ijk was a respondent to the first round of interviewing. In other words, the raked weights are adjusted by a factor equal to the ratio of the sum of the weights in the sample in the cell to the sum of the weights among only the respondents in the adjustment cell. The resulting weights are the initial nonresponse-adjusted weights for the 2015 panel.

2015 Survey File Weights

The 2015 Survey File data were collected in Fall 2015 from beneficiaries sampled in the 2012 through 2015 annual panels. To facilitate estimation from the resulting data, five sets of full-sample and replicate weights are derived. These include the 2015 continuously-enrolled cross-sectional weights; the 1-year longitudinal weights for analysis of 2014-2015 data,[36] 2-year longitudinal weights for analysis of 2013-2015 data, and 3-year longitudinal weights for analysis of 2012-2015 data; and finally, the 2015 ever-enrolled weights. In addition to the weights, the dataset includes the panel (selection year) identifier, and variance strata and variance unit variables for variance estimation. These variance strata and variance unit variables, along with the weights, capture all of the sampling design information necessary to estimate variances and make inferences to the population of Medicare beneficiaries.

Composition of Sample and Populations of Interest

The weights file includes records for beneficiaries that were sampled in the 2012, 2013, 2014, and 2015 Panels. The 2012, 2013, and 2014 panels are referred to as "continuing panels," while the 2015 Panel is referred to as the "incoming panel" as members of this sample were interviewed for the first time in Fall 2015. The Survey File weights include both continuously-enrolled and ever-enrolled weights in addition to the longitudinal weights. The continuously-enrolled weights represent a population of beneficiaries who were enrolled continuously between January 1st of the data year and completion of the Fall interview. The ever-enrolled weights represent the population of beneficiaries who were ever enrolled in Medicare for at least one day at any time during the data year.

The 2015 Survey File continuously-enrolled cross-sectional weights are populated for the subset of records with a completed Fall 2015 interview that are alive and entitled at the time of the interview. The resulting cross-sectional weights represent the population of beneficiaries that were continuously enrolled in Medicare from January 1, 2015, through completion of the Fall 2015 interview.

The one-year longitudinal weights are populated for members of the 2012, 2013, and 2014 panels that were continuously enrolled in both 2014 and 2015. The resulting weights represent the population of Medicare beneficiaries that enrolled on or before January 1, 2014, and are still alive and entitled as of completion of the Fall 2015 interview. The two-year longitudinal weights are populated only for members of the 2012 and 2013 panels who were continuously enrolled in each of the years 2013, 2014, and 2015. The population represented by these weights is the population of beneficiaries enrolled on or before January 1, 2013, and surviving and entitled as of completion of the Fall 2015 interview. Finally, the three-year longitudinal weights are populated only for members of the 2012 panel that were continuously enrolled during all of the years 2012-2015.

---

[36] The 2014 Survey File data were not released; as a result the 1-year longitudinal weight cannot be used for the 2015 data year.

The resulting weights represent the population of Medicare beneficiaries that enrolled on or before January 1, 2012, and are still alive and entitled as of completion of the Fall 2015 interview.

The 2015 Survey File ever-enrolled weights are populated for all records on the delivered file and include continuously enrolled beneficiaries and beneficiaries who died or lost entitlement prior to completing the Fall 2015 interview. Beneficiaries who first became enrolled in 2015 are also included; these current-year enrollees were sampled and interviewed for the first time in 2015. The resulting weights represent the population of beneficiaries that were enrolled in Medicare on at least one day at any point in 2015.

Continuously-Enrolled Cross-Sectional Weights

The continuously-enrolled cross-sectional weights are the traditional Survey File weights and have been provided every year. They represent the population of beneficiaries that were enrolled in Medicare for the entire period between the first of the year through the Fall 2015 interview period.

Fall 2015 Nonresponse Adjustment

Continuing sample from the 2012, 2013, and 2014 panels are adjusted for nonresponse through Fall 2015. The process begins with weights for these panels that were previously adjusted through Fall 2014. Response status in Winter/Summer 2015 and Fall 2015 is then identified, where a respondent is a beneficiary that was alive and entitled with a complete Fall 2015 interview, or who died or lost entitlement at some time in Winter/Summer 2015 and Fall 2015 but had a completed final interview after death (via proxy) or loss of entitlement.

Nonresponse adjustment cells are constructed prior to performing the adjustment. First, the beneficiaries are divided into five primary adjustment cells: alive community, deceased community, alive facility, deceased facility, and Fall 2014 nonrespondents.

Separately within each of these main adjustment cells, and separately by panel, response propensity models are fit using logistic regression to model the probability of response through Fall 2015 as a function of covariates derived from the Fall 2014 ATC data. Generally, the covariates are selected into the logistic regression model using stepwise selection with an entry p-value of .10 and a stay p-value of .15. Using the predicted response probabilities, beneficiaries are grouped into cells of approximately 100 each. Across all panels there are a total of 118 adjustment cells formed following the response modeling process. Separately within each of these cells, a ratio adjustment to distribute the weights of the nonrespondents to the respondents is performed. The resulting weights are the within-panel weights adjusted for response through Fall 2015.

Derivation of the Continuously-Enrolled Weights

The next step takes the weights for continuing panels that are now adjusted through Fall 2015 and combines them with the weights for the 2015 Panel that were separately adjusted for initial nonresponse at the first interview (Fall 2015) as part of the initial weighting process. Next, the process removed cases that either died or lost entitlement prior to the Fall 2015 interview, or were cases from the 2015 panel that enrolled after January 1, 2015.

At this stage there is quadruple coverage of beneficiaries who accreted on or before January 1, 2012, triple coverage of beneficiaries who accreted from January 2, 2012 through January 1, 2013, and double coverage of beneficiaries who accreted from January 2, 2013 through January 1, 2014. To account for this overlap, the weights for the four panels are adjusted by compositing factors derived from the number of effective completes by accretion year and age group across the four panels.

The compositing factor applied to beneficiaries from panel p in accretion year/age group domain d is

$$\varphi_{pd} = \frac{n_{pd}^{eff}}{\sum_{p \in P} n_{pd}^{eff}}$$

where n_{pd}^{eff} is the effective number of Fall 2015 completes in panel i in accretion year/age group domain d. The subscript p indexes the four panels in the set of active panels P. The effective sample sizes are calculated as

$$n_{pd}^{eff} = \frac{n_{pd}^{act}}{1 + \left(\frac{s_{pd}^{73}}{\overline{w}_{pd}^{73}} \right)^2}$$

where n_{pd}^{act} is the actual number of completed interviews, $\overline{w}_{pd}^{73}$ is the average of the Fall 2015 adjusted weights for the panel, and s_{pd}^{73} is the standard deviation of these weights.

The resulting weights are the final continuously-enrolled cross-sectional weights for the 2015 Survey File (SF). They represent the 2015 continuously-enrolled population.

Longitudinal Weights

The derivation of one-year longitudinal weights begins with the weights adjusted through Fall 2015 for these panels as described above, subset to beneficiaries that were alive and entitled at the Fall 2015 interview. The weights were then further adjusted to account for triple coverage of those accreting on or before January 1, 2012, and double coverage of those accreting from January 2, 2012 through January 1, 2013, using compositing factors derived similarly as described in the previous section. The final resulting weights represent the one-year longitudinal population, which is the population of beneficiaries that enrolled on or before January 1, 2014, and were alive and entitled as of the Fall 2015 interview.

The derivation of two-year longitudinal weights begins with the weights adjusted through Fall 2015 for these panels, subset to beneficiaries that were alive and entitled at the Fall 2015 interview. The weights are then further adjusted to account for double coverage of those accreting on or before January 1, 2012, using compositing factors. The final resulting weights represent the two-year longitudinal population, which is the population of beneficiaries that enrolled on or before January 1, 2013, and were alive and entitled as of the Fall 2015 interview.

The three-year longitudinal weights are comprised of members of the 2012 panel and are equal to the weights adjusted through Fall 2015 for this panel, subset to beneficiaries that were alive and entitled at the Fall 2015 interview. There is no need for further adjustment by compositing factors because there is only one panel providing three-year data so the weights are equal to the final cross-sectional weights for these beneficiaries. The final weights represent the three-year longitudinal population, which is the population of beneficiaries that enrolled on or before January 1, 2012, and were alive and entitled as of the Fall 2015 interview.

Final Ever-Enrolled Weights

Ever-enrolled Survey File weights were introduced for the first time in 2015. These weights represent the population of Medicare beneficiaries who were ever enrolled at any time during 2015 (i.e., enrolled on at least one day in 2015). The continuously-enrolled beneficiaries are a subset of the ever-enrolled beneficiaries in two ways, both in terms of the real-world populations they represent and in terms of the sampled and interviewed beneficiaries that appear on the Survey File.

Fall 2015 Nonresponse Adjustment

Continuing sample from the 2012, 2013, and 2014 panels are adjusted for nonresponse through Fall 2015. As with the continuously-enrolled weights, the process begins with weights for these panels that were previously adjusted through Fall 2014. The response status in Winter/Summer 2015 and Fall 2015 is then identified. Under the ever-enrolled design, respondents include beneficiaries with a complete Fall 2015 interview, those who lost entitlement prior to Fall 2015 and had a final complete interview, those who died prior to Fall 2015 whether or not a final proxy interview was obtained, and Fall 2015 nonrespondents who were successfully re-fielded in Winter 2016.

Next, the weights are adjusted for nonresponse through Fall 2015, using the same cells that are created for the adjustment of the weights under the continuously-enrolled design. Following ratio adjustments within these cells, the resulting weights are the within-panel weights adjusted for response through Fall 2015 for purposes of the ever-enrolled weights.

Derivation of the Ever-Enrolled Weights

The next step begins with the weights for the continuing panels adjusted through Fall 2015 in the previous step and combines them with the weights for the 2015 Panel that are separately adjusted for initial nonresponse at the first interview (Fall 2015). Next, the small number of cases that died or lost entitlement prior to January 1, 2015, and hence were never enrolled in 2015, are removed.

At this stage, beneficiaries from the continuing panels who died or lost entitlement during 2015 are included, as described in the previous section. However, the 2015 panel cases include only those who were respondents to the Fall 2015 initial interview, and as such they do not include any beneficiaries that died or lost entitlement prior to Fall 2015. Beneficiaries who enrolled on or before January 1, 2014, who died or lost entitlement are accounted for by the continuing panels. Enrollees after January 1, 2014, who died or lost entitlement are not represented by any other panels, but they are few in number and are accounted for during final poststratification.

As with the continuously-enrolled and longitudinal weights, the ever-enrolled weights for the four panels are adjusted by compositing factors to account for overlap between the panels. These are derived from the number of effective completes by accretion year and age group. For the ever-enrolled weights, beneficiaries from the continuing panels who died or lost entitlement in 2015 are combined separately to account for the fact that these beneficiaries are not represented by the 2015 Panel.

To finalize the ever-enrolled weights, the raking technique to calibrate the weights to known population control totals for the ever-enrolled population is used. These are derived from the enrollment data extracts for drawing the 2015 Panel. The raking dimensions used are age category (7-level) and accretion year (6-level). The raking process adjusts the weights to match the control totals for the first raking dimension, then for the second raking dimension, then for the first dimension again, and so on until the weights perfectly match the control totals in both dimensions. The resulting weights are the final ever-enrolled weights for 2015. They represent the population of beneficiaries that were enrolled for at least one day at any time in 2015. Exhibit 8.3.2 and 8.3.3 present the control totals used for the raking adjustment step.

Exhibit 8.2.2: Control Totals for Ever-Enrolled Weight Raking, Dimension 1: Age Group

Age Group	Control Total
< 45 Years	1,918,920
45 -64 Years	7,160,100
65 - 69 Years	15,550,740
70 - 74 Years	11,537,220
75 - 79 Years	8,224,140
80 - 84 Years	5,930,300
85+ Years	6,954,560
Total	**57,275,980**

Exhibit 8.2.3: Control Totals for Ever-Enrolled Weight Raking, Dimension 2: Enrollment Year

Enrollment Year	Control Total
< 2011	39,574,860
2011	3,439,620
2012	3,781,980
2013	3,653,580
2014	3,544,900
2015	3,281,040
Total	**57,275,980**

2015 Cost Supplement Weights

Data for the 2015 Cost Supplement File were collected in Winter 2015 through Winter 2016. The weights include beneficiaries sampled in the 2012 through 2014 Panels, plus members of the 2015 Panel who were recently enrolled in Medicare. These Cost Supplement File weights are ever-enrolled weights representing the population of beneficiaries who were enrolled for at least one day in 2015. In addition to the weights, the dataset includes panel (selection year) identifier, and variance strata and unit variables for variance estimation.

Composition of Sample and Populations of Interest

The 2015 Cost Supplement weights include beneficiaries that were sampled in the 2012, 2013, 2014, and 2015 Panels. The 2012, 2013, and 2014 Panels are referred to as "continuing panels" and provide survey-reported cost and utilization for 2015 through participation in the MCBS during Winter 2015 through Winter 2016. Members of the 2015 Panel who were first enrolled in 2014 (excluding January 1, 2014) or 2015 are referred to as "recent enrollees." They were first interviewed in Fall 2015 and did not provide cost and utilization data for the period of time between enrollment and completion of the Fall 2015 interview; cost and utilization data for the period between the Fall 2015 interview and the end of 2015 were collected in Winter 2016. A combination of the survey-collected data for the end of year and Medicare claims data will be used to impute beneficiary-level data for the entire period of enrollment in 2015. The final weights, which include both the continuing panels and the recent enrollees, represent the population of beneficiaries that were ever enrolled in Medicare at any time during 2015.

Adjustment Derivation of Cross-Sectional Weights for the Continuing Panels

The process begins with weights for the 2012, 2013, and 2014 panels that were previously adjusted through Fall 2015 as part of the 2015 Survey File weights. These weights are further adjusted based on a product of the 2015 reference period process that identifies which beneficiaries contributed enough cost and utilization data to be included in the final data products. To be included, sample members must meet at least one of the following three criteria: (a) the ratio of days covered by interviews to the number of days enrolled in Medicare in 2015 is equal to or greater than 0.66; (b) the difference between the number of days enrolled in Medicare and the number of days covered by interviews is less than or equal to 60 days; or (c) the beneficiary is a recent enrollee from the 2015 Panel who completed the initial Fall 2015 interview. Beneficiaries that died or lost entitlement prior to January 1, 2015, are ineligible and removed at this stage. Beneficiaries who survived into 2015 but do not meet the above criteria are considered to be nonrespondents for the 2015 Cost Supplement File and are adjusted for in the resulting weights. The adjustment cells used for this ratio adjustment are the same cells that were created during weighting for the 2015 Survey File weights.

Note that at this stage there is triple coverage of beneficiaries who accreted on or before January 1, 2012, in the continuing panels, and double coverage of beneficiaries who accreted from January 2, 2012 through January 1, 2013. Therefore the weights for the three panels are adjusted by compositing factors derived from the effective number of completes by panel, accretion year, and age group. The resulting weights are the pre-raked cross-sectional weights for the continuing panels.

Cross-Sectional Weights for the Recent Enrollees

The "recent enrollees" are those who enrolled between January 2, 2014, and December 31, 2015, inclusive. This step begins with the initial weights for the 2015 Panel, adjusted for nonresponse at the Fall 2015 interview. The subset of all Fall 2015 respondents from the 2015 Panel that are recent enrollees is isolated, and the resulting weights for this subset are the pre-raked cross-sectional weights for the recent enrollees.

Final Cross-Sectional Ever-Enrolled Weights for the Cost Supplement

The sum of the combined weights across all four panels (the three continuing panels plus the recent enrollees from the 2015 Panel), provides an estimate of the ever-enrolled population in 2015, but is not exact. To finalize the ever-enrolled weights, the raking technique is used to calibrate the weights to known population control totals for the ever-enrolled population. The raking dimensions used are age category (7-level) and accretion year (6-level), and the control totals used are the same as those used for the Survey File ever-enrolled weights calibration presented in Exhibit 8.3.2. The resulting weights are the final weights for the 2015 Cost Supplement File. They represent the population of beneficiaries that were enrolled for at least one day at any time in 2015.

2015 Topical Module Weights

The Beneficiary Knowledge Questionnaire (KNQ) was administered in the community questionnaire in Winter 2016 (Round 74). The Summer 2016 (Round 75) community questionnaire included the Income and Assets Questionnaire (IAQ), the Patient Activation Questionnaire (PAQ) and the Prescription Medicine Questionnaire (RXQ). To facilitate estimation from the resulting data, two sets of full-sample and replicate weights were derived for each module, one based on the 2015 Survey File continuously-enrolled population, and the other based on the 2015 Cost Supplement ever-enrolled population. These weights can be used to conduct joint analyses of Topical Module data, Survey File data, and Cost Supplement data. Exhibit 8.3.4 lists the Topical Module weights for these rounds.

Exhibit 8.2.4. 2015 Data Year Topical Module Survey Weights Datasets and Contents

Dataset Name	Record Count	Variable Count	Full-Sample Weight	Replicate Weights	Description
KNCWT	10,375	102	KN74CWT	KNC1-KNC100	R74 KNQ Continuously-Enrolled
KNEWT	7,801	102	KN74EWT	KNE1-KNE100	R74 KNQ Ever-Enrolled
IACWT	7,266	102	IA75CWT	IAC1-IAC100	R75 IAQ Continuously-Enrolled
IAEWT	5,351	102	IA75EWT	IAE1-IAE100	R75 IAQ Ever-Enrolled
PACWT	6,438	102	PA75CWT	PAC1-PAC100	R75 PAQ Continuously-Enrolled
PAEWT	4,711	102	PA75EWT	PAE1-PAE100	R75 PAQ Ever-Enrolled
RXCWT	4,993	102	RX75CWT	RXC1-RXC100	R75 RXQ Continuously-Enrolled
RXEWT	3,178	102	RX75EWT	RXE1-RXE100	R75 RXQ Ever-Enrolled

Composition of Sample and Populations of Interest

The topical module data were collected from beneficiaries selected in the 2012, 2013, 2014, or 2015 panels who responded to the community questionnaire in the round in which the module was administered, and each of the datasets includes members of these panels who completed the topical module. Each of the eight weights were derived to represent a population that was alive, entitled, and in the community in the given round.

The four weights that began with the 2015 Survey File continuously-enrolled weights as a starting point (KNCWT, IACWT, PACWT, and RX7WT) all represent populations that were continuously enrolled from the start of 2015 and were still alive, entitled, and in the community during the round. The KNCWT weights represent the population of beneficiaries that were continuously enrolled from January 1, 2015 and still alive, entitled, and in the community in the Winter of 2016. The IACWT, PACWT, and RXCWT weights represent the population of beneficiaries that were continuously enrolled from January 1, 2015 and still alive, entitled, and in the community in Summer 2016.

The four weights that began with the 2015 Cost Supplement ever-enrolled weights as a starting point (KNEWT, IAEWT, PAEWT, and RXEWT) all repre sent populations that were ever enrolled at any time in 2015 and were still alive, entitled, and in the community during the round. The KNEWT weights represent the population of beneficiaries that were ever enrolled for at least one day in 2015 and were still alive, entitled, and in the community in the Winter of 2016. The IAEWT, PAEWT, and RXEWT weights represent the population of beneficiaries that were ever enrolled for at least one day in 2015 and were still alive, entitled, and in the community in the Summer of 2016.

Derivation of Topical Module Weights

Each of the topical module weights is based on a starting weight, which is either the 2015 Survey File continuously-enrolled weight (those that end in 'CWT' or the 2015 Cost Supplement ever-enrolled weight (those that end in 'EWT'). The choice of starting weight determines the population that the derived topical module weight represents, as described in the previous section. However, once this choice of starting weight is made, the process for each topical module is largely the same.

The weighting adjustments for each delivered weight are carried out in two steps. At each, the existing model-based adjustment cells that were developed for the 2015 Survey File and Cost Supplement weights were used, with collapsing of the cells where necessary to preserve adequate sample sizes.

The first adjustment distributes the weights for cases with unknown eligibility for the module to those with known eligibility. Beneficiaries may have unknown eligibility if they were unlocatable during the round or if they were nonrespondents during the round or earlier rounds and there was no indication of mortality or residential (community or facility) status. The number of cases with unknown eligibility was small in Winter 2016 because this round immediately followed 2015 and the Fall 2015 Survey File interviews, whereas in Summer 2016 there was an intervening round in which some members of the sample became nonrespondents. In all cases, this first adjustment for unknown eligibility makes the implicit assumption that if eligibility were observable for these cases they would exhibit the same proportions of eligibility as the cases whose eligibility we are able to observe.

Prior to the second adjustment, the set of beneficiaries is limited to those who were eligible to receive the topical module. A beneficiary was considered ineligible if they had died, lost entitlement, or were in the Facility component only during the round. The second adjustment, the nonresponse adjustment, then distributes the weights for the eligible nonrespondents to the eligible respondents. For purposes of the PAQ, alive and entitled beneficiaries in the community for whom the community questionnaire was administered by proxy were considered nonrespondents. For the RXQ, the exit panel (2012) was not administered the module. Therefore, the weights are limited only to beneficiaries from the 2013, 2014, and 2015 panels. To account for the loss of the 2012 panel, the nonresponse adjustment also included an additional adjustment to increase the weight totals so that they agree with the total weights for the IAQ and PAQ, which were conducted in the same quarter as the RXQ.

8.3 MCBS Imputation Processes

Overview

As noted earlier, MCBS imputation falls under two umbrellas that focus on imputing monetary amounts: Income and Asset (IA) imputation, and Event, Payer, and Cost imputation, which includes imputation for Prescription Medicine (PM) and Non Prescription Medicine (Non PM) events and costs. All three imputations focus on imputing a monetary amount. IA imputation completes income and asset information for the beneficiary and spouse, and PM and Non PM imputation complete medical event and cost data. For all three types, two groups of variables are imputed:

- Probes: Yes/no variables indicating whether the type of income, asset, or payer should have a nonzero amount.
- Amounts: The value of the income, asset, or cost paid for a medical event. For IA imputation, amounts are nonzero if the associated probe indicates the income or asset exists and missing otherwise. For PM and Non PM imputation, amounts are nonzero if the associated probe indicates that the payer paid and zero otherwise.

For both probes and amounts, single value imputation is performed sequentially from variables or records with the least to the most item nonresponse.

Income and Asset Imputation

Overview

The 2015 Income and Assets (IA) imputation imputes detailed information about income and assets of the beneficiary and spouse for Community Questionnaire respondents. For Facility Questionnaire respondents, and Community and Facility Questionnaire non-respondents[37], only total income was imputed due to the lack of detailed asset information.

Process

Respondents are asked about their prior year income and assets during the Summer round. The income and asset data first go through data editing to ensure that respondent-reported values were either appropriate or set to missing. Data editing is performed to:

- Match skip logic within the Income and Asset Questionnaire (IAQ)
- Set extreme outliers at the tails of the distributions of each IA variable to missing
- Set outliers based on joint distributions of highly-correlated IA variables to missing

---

[37] The Income and Assets questionnaire section (IAQ) is only administered once per year. Non-response to this section may be due to non-response in the round the questionnaire section was to be administered, or non-response to questions in the IAQ. For more information on IAQ, see Section 4.1.

Next, probe variables are imputed via a hot deck method. Probes had very low item nonresponse rates. The hot deck method was chosen because it could impute all of the missing values and was relatively easy to implement. This method takes the non-missing IA value directly from another beneficiary in the same imputation cell to fill in the missing IA value of the recipient beneficiary. If the probe is imputed as "no", indicating that a beneficiary does not have a particular type of asset, the corresponding amount variable is set to missing.

Amount variables are imputed after probes. While most respondents report whether the beneficiary has an asset type, some respondents refuse to provide or don't know the amount of the asset. As a result, amount variables need more imputation. When respondents report value ranges, the hot deck method is used to impute an exact dollar amount using the given value range as a boundary. When value ranges are not given but prior-year IA information exists, values are imputed using a prior-year carry-forward method with an inflation adjustment. This method uses the non-missing IA variable value for the same beneficiary and variable from the prior year to impute the current-year missing value. This prior-year carry-forward method provides reliable imputed values for these respondents. For the rest of the missing amount values, hot deck imputation is used.

Each variable imputed via hot deck imputation has a unique set of imputation cell variables. In the hot deck method, recipient and donor records are segregated into pools of records ("imputation cells") that have the same values on a set of auxiliary (or explanatory) variables. In general, the auxiliary variables that define imputation cells for probe variables include prior-year probe values, beneficiary's age, indicator of spouse/partner, and other related IA probes. Auxiliary variables that define imputation cells for amount variables include other related IA amounts, poverty indicators, beneficiary's age, and metropolitan status.

Prescription Medicine and Non Prescription Medicine Imputation

Overview

Both the Prescription Medicine (PM) imputation and Non Prescription Medicine (Non PM) imputation fill in missing payer and payment information for beneficiaries' medical events. The imputation procedures used for Prescription Medicine (PM) events versus all other event types (Non PM) are very similar but not identical.

Process

Both PM and Non PM imputation begin with the receipt of the survey-reported events matched against the Medicare claims. Three categories of records are returned: events found in the claims only (claims-only), events found in the survey-reported data only (survey-only), and survey-reported events that were successfully matched to a Medicare claim (survey-matched).

For the PM imputation, only unmatched survey-only events are processed through imputation. Claims-only and survey-matched events are considered complete. For the Non PM imputation, all three claims match statuses are processed through imputation.

First, data preprocessing and editing are performed to identify the total payment for the event and the most likely payers for the event. This procedure is described in detail in the MCBS Data User's Guide: Cost Supplement File. Imputation then proceeds in three steps.

First imputed are events where the total payment is known and the payers and payment amounts are missing together (when a payer is missing, the amount is missing, and vice versa). Exhibit 8.3.1 gives an illustration of the type of record that would be imputed in this group, with a simplified potential payer vector. The donor record is required to be a complete record, and must have at least one of the recipient's missing payers as a payer with a positive payment amount, so that there is at least one amount value to which the difference between the total payment and the sum of the known payments can be allocated. In the example shown in Exhibit 8.3.1, a donor would need to have either "Employment-based private health insurance" or "Out of Pocket" as a payer with a nonzero amount. The payers and payment amounts are pulled from the same donor.

Exhibit 8.3.1. Payers and Payment Amounts Missing Together, Total Payment Known

Variable Type	Medicare Fee-for-Service	Medicaid	Employment-based private health insurance	Out of Pocket	Total Payment
Payer Indicator	Yes	No	(null)	(null)	--
Amount	50	0	(null)	(null)	200

Next imputed are events where the total payment is known and the payers and payment amounts have different missing patterns (i.e., there is at least one instance where the payer is known to have paid but the amount is missing). This is illustrated by Exhibit 8.3.2. The payers are imputed first. Donors are required to be complete records. There is no restriction that the donor is a payer for any of the the recipient's missing payers because by definition of this group, there is at least one known payer already to which the missing payment amount can be allocated. Payment amounts are imputed next. If the payer is imputed not to have paid, the payment amount is set to zero. If there is only one missing payment amount after the payer imputation, that amount is completed by subtraction. If possible, payment amounts are all pulled from the same donor; if a donor with the required payer pattern does not exist[38], payment amounts are imputed individually from different donors.

---

[38] In this group, we impute a vector of missing payers together from the same donor, and have at least one additional payer who is known to have paid but the amount is unknown. Thus, a new payer pattern that did not exist in the original data may be created – the vector of imputed payers, plus the known payer with unknown amount.

Exhibit 8.3.2. Payers and Payment Amounts Missing Differentially, Total Payment Known

Variable Type	Medicare Fee-for-Service	Medicaid	Employment-based private health insurance	Out of Pocket	Total Payment
Payer Indicator	Yes	No	Yes	(null)	--
Amount	50	0	(null)	(null)	(null)

Finally, events with the total payment unknown are imputed (illustrated by Exhibit 8.3.3). Payers are imputed first and are all taken from the same donor. Payment amounts are imputed next and are taken from the same donor when possible, or are imputed individually if a donor with the required payer pattern does not exist[39]. Total payment is set to the sum of the payment amounts.

Exhibit 8.3.3. Total Payment Unknown

Variable Type	Medicare Fee-for-Service	Medicaid	Employment-based private health insurance	Out of Pocket	Total Payment
Payer Indicator	Yes	No	Yes	(null)	--
Amount	50	0	(null)	(null)	(null)

In all PM and nearly all Non PM cases, the payment amount is not imputed directly from the donor; it is ratio-adjusted to fit with the recipient's known payment amounts.

The PM and Non PM imputation processes are very similar up to this point but then diverge.

PM Imputation

One final step is applied in PM imputation processing. After the general imputation procedure has been run, cases are reviewed and those found to be inconsistent or to have potential imputation issues are reviewed and edited. Records where the payers and payment amount vectors are complete but total payment is less than or more than the sum of the payment amounts, or records that are incomplete but have total payment less than the known payment amounts, are subjected to edits to make the record complete and consistent. Events where an imputed payment amount is less than a penny or a total payment is less than 50 cents are re-imputed

---

[39] Similar to when total payment is known, some records with total payment unknown will have payers and payment amounts missing at different rates (i.e., there is at least one instance where the payer is known to have paid but the amount is missing). After the payer imputation, a new payer pattern may be created that did not exist in the original data.

from a new donor. The number of records requiring editing or re-imputation is very small (1.5% of records in 2015).

The PM imputation produces one file, an event-level dataset of survey-only events.

Non PM Imputation

New in 2015, current-year enrollee sample beneficiaries are included in the Non PM imputation. [40] The current-year enrollees have some portion of the year covered by claims data only, and not by survey data. This may result in biased estimates as some medical events and costs, such as vision and dental health care services, are not covered by the Medicare claims and would be captured only by the survey data that were not collected. Please see the MCBS Data User's Guide: Cost Supplement File for a further discussion of gaps in survey data coverage. A new unit-level imputation procedure was added to address the issue of gaps in survey data coverage for the current-year enrollees. This procedure imputes survey-only events that may not be covered by the claims, adding new event records to the file that did not previously exist.

The time period within which claims-only events are to be imputed varies by individual, ranging from the beneficiary's enrollment date to the first of: the fall interview date (if there was a completed winter interview), the date of death, the date of lost entitlement, or December 31. First, this time period (the "Missing Period") is defined for each current-year enrollee. A donor is selected for each current-year enrollee, and the donor's survey-only records (excluding those with a Medicare and not Medicare Advantage payment, as these would be covered by claim data) that occur within the recipient's Missing Period are then created for the recipient. If the donor has no donation-eligible records of a given event type, no records are created.

All variables populated on the donor record are populated on the newly-created (recipient) record. Variables that relate to the event are pulled along from the donor record. Variables that relate to the beneficiary are retained from the recipient.

As described in the MCBS Data User's Guide: Cost Supplement File, the event types used in the survey differ from the event types in the Medicare claims. For the Non PM events, an administrative event type is imputed from the survey-reported event type. Event type imputation recipients are events found in the survey-only data, and donors are survey-matched events. Recipient records are matched to donors on survey-reported event type and cost, and the donor's administrative event type is assigned to the recipient.

Next, hospice event data are appended to the Non PM events. These data come directly from CMS and are not imputed. More information on hospice data is provided in the MCBS Data User's Guide: Cost Supplement File.

Finally, the Non PM data are aggregated to the service and person level. The Non PM imputation produces three files: at the event level (most disaggregate), at the service level (one record per beneficiary and event type), and at the person level (one record per beneficiary). Event-level

---

[40] See Section 3.4, "Current-Year Enrollee Sample", for more information these beneficiaries.

records are first summed to the service level, and then adjustments are performed to annualize these amounts and adjust for days the beneficiary was eligible for Medicare but not covered by survey-reported data. This process is described in further detail in the MCBS Data User's Guide: Cost Supplement File. Then, unadjusted and adjusted service-level amounts are summed to the person level.

Hot Deck Imputation Procedure

All PM and Non PM imputation is performed using a hot deck imputation procedure.

While hot deck has been used as a donor selection method for several years on the MCBS, the method to identify a compatible donor is new for 2015.

Each imputation step has a unique set of qualification rules and key variables used to identify a similar donor record for a given recipient record. The donor pool for each set of recipients is first restricted to the group of potential donor records that meets the donor qualification rules, such as requiring that donors have complete data on the item to be imputed. Next, the similarity between a given recipient and each possible donor is measured via the Gower function using SAS/STAT® software's PROC DISTANCE:

$$s_1(x, y) = \frac{\sum_{j=1}^{v} w_j \delta_{x,y}^{j} d_{x,y}^{j}}{\sum_{j=1}^{v} w_j \delta_{x,y}^{j}}$$

Where v is the number of variables, x_j is the data for observation x and the j^{th} variable, y_j is the data for observation y and the j^{th} variable, and w_j is the weight for the j^{th} variable. For ordinal, interval, and symmetric nominal variables, $\delta_{x,y}^{j} = 1$. For asymmetric nominal variables, $\delta_{x,y}^{j} = 1$ if either x_j or y_j is present and 0 if both are absent. For a nominal variable, $d_{x,y}^{j} = 1$ if $x_j = y_j$ and 0 otherwise. For an ordinal, interval, or ratio variable, $d_{x,y}^{j} = 1 - |x_j - y_j|$.[41,42,43]

The Gower function was selected because it can compute a similarity measure across several variable types (nominal, ordinal, and interval). For each recipient, we select donors whose similarity score is less than or equal to the 30th largest distance (with a score of 0 representing identical records and 1 representing divergent records). This may result in 30 potential donors, or more if there are ties. Frequently, PM and Non PM donor pools are small, and this method allows us to relax some of the boundaries defining a suitable donor while continuing to find donors that are highly similar to a recipient. After computing donor pools by finding donor records that are similar to recipients, the new imputation procedure goes on to identify the donor record using the hot deck method in SAS/STAT software's PROC SURVEYIMPUTE.

---

[41] SAS Institute Inc. 2017. SAS/STAT® 14.3 User's Guide. Cary, NC: SAS Institute Inc.

[42] Podani, J. (1999). Extending Gower's General Coefficient of Similarity to Ordinal Characters. *Taxon, 48*(2), 331-340.

[43] Gower, J. (1971). A General Coefficient of Similarity and Some of Its Properties. *Biometrics, 27*(4), 857-871.

9. RESPONSE RATES AND NONRESPONSE

This section presents the response rates and describes the derivation of those rates for the 2015 Cost Supplement and Survey File data releases.

9.1 Response Rates

This section details the definitions and calculations of Cost Supplement File response rates and Survey File response rates. Response rates presented in this report are unweighted.

In the sections that follow, both unconditional and conditional response rates are presented. The unconditional response rate is the percentage of sample that were released during the fall round of the selection year and responded to the survey in 2015. The unconditional response rates, also called cumulative response rates, use the original selected sample size as the baseline in their calculation. Conditional response rates are the percentage of sample that responded during 2014 and also responded during 2015. Conditional response rates use the sample who responded during 2014 as the baseline in their calculation. In other words, they are conditioned on response in year 2014.

2015 Cost Supplement File Response Rates

Unconditional Response Rates for the Annual Cost Supplement File

The response rate for a given data year, t, in canonical form is simply

$$r_t = \frac{C_t}{E_t},$$

where C_t is the number of beneficiaries for whom the Cost Supplement File data are taken to be *complete*, and E_t is the number of beneficiaries who are considered *eligible* for the annual Cost Supplement File data release.

C_t is calculated as the number of beneficiaries with a non-missing, positive Cost Supplement File weight for the given year.

The number of eligible beneficiaries is calculated as

$$E_t = T_t - I_t ,$$

where T_t is the *total sample size* for the given year, and I_t is the number of beneficiaries who are considered *ineligible* for the given annual Cost Supplement File data release.

For the $t = 2015$ data year, T_t includes the following:

- All of the panel selected in year $t - 3$, called S_{t-3}.

- All of the panel selected in year $t - 2$, called S_{t-2}.
- All of the panel selected in year $t - 1$, called S_{t-1}.
- The subset of the panel selected in year t, called S_t, consisting of members of both the year $t - 1$ and the year t cohorts of beneficiaries.

Conditional Response Rates for the Annual Cost Supplement File

The conditional response rate for the year $t - 3$ to $t - 1$ panels in Cost Supplement File year t is:

$$\frac{C_t}{E_t - N_t},$$

where

$C_t = S_{t-3}$ to S_{t-1} panel beneficiaries with positive weights on the year t Cost Supplement File;

$E_t = S_{t-3}$ to S_{t-1} panel beneficiaries still entitled on January 1, year t;

$N_t =$ subset of E_t that were not released in the first round of year t.

The conditional response rate for the year t panel in Cost Supplement File year t is:

$$\frac{C_t}{E_t},$$

where

$C_t = S_t$ panel beneficiaries with positive weights on the Cost Supplement File;

$E_t = S_t$ panel beneficiaries enrolled between January 2, year $t - 1$ to December 31, year $t - 1$ and still entitled on January 1, year t.

The conditional response rate for the year $t + 1$ panel in Cost Supplement File year t is:

$$\frac{C_t}{E_t},$$

where

$C_t = S_{t+1}$ panel beneficiaries with positive weights on the Cost Supplement File;

$E_t = S_{t+1}$ panel beneficiaries enrolled between January 1, year t and December 31, year t.

Exhibits 9.1.1 and 9.1.2 display the 2015 Cost Supplement File unconditional and conditional response rates by panel.

Exhibit 9.1.1: 2015 MCBS Annual Cost Supplement File Unconditional Response Rates

Panel	Released	Complete	Eligible	Ineligible	Unconditional Response Rate
2012	7,400	2,356	6,358	1,042	37.1%
2013	7,400	2,693	6,679	721	40.3%
2014	11,398	4,578	10,766	632	42.5%
2015	701	354	694	7	51.0%
Total	26,899	9,981	24,497	2,402	40.7%

Exhibit 9.1.2: 2015 MCBS Annual Cost Supplement File Conditional Response Rates

Panel	Complete	Eligible	Subset of Eligibles Not Released	Conditional Response Rate
2012	2,356	6,358	3,208	74.8%
2013	2,693	6,679	2,771	68.9%
2014	4,578	10,766	4,452	72.5%
2015	354	694	0	51.0%
Total	9,981	24,497	10,431	71.0%

2015 Survey File Response Rates

Unconditional Response Rates for the Annual Survey File: Ever-Enrolled Beneficiaries.

The response rate for a given data year, t, in canonical form is simply

$$r_t = \frac{C_t}{E_t},$$

where C_t is the number of beneficiaries for whom the Survey File data are taken to be *complete*, and E_t is the number of beneficiaries who are considered *eligible* for the annual Survey File data release.

C_t is calculated as the number of beneficiaries with a non-missing, positive Survey File ever-enrolled weight for the given year.

The number of eligible beneficiaries is calculated as

$$E_t = T_t - I_t ,$$

where T_t is the *total sample size* for the given year and I_t is the number of beneficiaries who are considered *ineligible* for the given annual Survey File data release.

For year t, T_t includes the following:

- All of the panel selected in year $t-3$, called S_{t-3}.
- All of the panel selected in year $t-2$, called S_{t-2}.
- All of the panel selected in year $t-1$, called S_{t-1}.
- All of the panel selected in year t, called S_t .

I_t is calculated as the number of beneficiaries from panels $t-3$ to $t-1$ who died or lost entitlement prior to January 1st of year t, plus the number of ineligible or deceased beneficiaries from the year t panel in the fall round.

Conditional Response Rates for the Annual Survey File: Ever-Enrolled Beneficiaries.

The conditional response rate for the year $t-3$ to $t-1$ panels in Survey File year t is:

$$\frac{C_t}{E_t - N_t},$$

where

$C_t = S_{t-3}$ to S_{t-1} panel beneficiaries with positive weights on the year t Survey File;

$E_t = S_{t-3}$ to S_{t-1} panel beneficiaries still entitled and alive prior to fall round, year t and are not I_t.

$N_t = \text{subset of } E_t$ that were not released in the first round of year t.

The conditional response rate for the year t panel in Survey File year t is:

$$\frac{C_t}{E_t},$$

where

$C_t = S_t$ panel beneficiaries with positive weights on the Survey File;

$E_t = S_t$ panel beneficiaries still entitled and alive prior to fall round, year t and are not I_t.

Response Rates for the Annual Survey File: Continuously-Enrolled Beneficiaries

The formulas for calculating the unconditional and conditional response rates for the continuously-enrolled beneficiaries are identical to the corresponding formulas detailed above for the ever-enrolled population. The only differences are in the definitions of C_t and I_t .

For the continuously-enrolled response rate calculations, C_t is calculated as the number of beneficiaries completing an interview in the fall round of year t with a non-missing, positive Survey File continuously-enrolled weight for the given year t.

Two subsets of ineligibles contribute to I_t for the continuously-enrolled response rate calculations:

- The first subset includes beneficiaries who are found to be ineligible or deceased in any round up to and including the fall round of year t.
- The second subset includes beneficiaries who finished the fall round year t interview but are not Survey File completes, *or* beneficiaries who were non-respondents prior to the fall round of year t and thus were not fielded in the fall round, and had a final status with no further attempts to field in any previous round. (These are beneficiaries not included in the first subset of ineligibles described above.) For these cases, the date of death or lost entitlement date, if any, is compared to the average interview date in the fall round year t. If date of death or lost entitlement date is prior to the average interview date, the case is determined to be ineligible. Otherwise, it is determined to be an eligible non-respondent.

Exhibits 9.1.3 and 9.1.4 display the 2015 annual Survey File unconditional response rates by panel for ever-enrolled and continuously-enrolled beneficiaries.

Exhibit 9.1.3: 2015 MCBS Annual Survey File Unconditional Response Rates for Ever-Enrolled Beneficiaries

Panel	Released	Ever-Enrolled Complete	Ever-Enrolled Eligible	Ever-Enrolled Ineligible	Unconditional Response Rate of Ever-Enrolled Beneficiaries
2012	7,400	2,380	7,298	102	32.6%
2013	7,400	2,716	7,300	100	37.2%
2014	11,398	4,626	11,255	143	41.1%
2015	8,621	4,349	8,179	442	53.2%
Total	34,819	14,071	34,032	787	41.3%

Exhibit 9.1.4: 2015 MCBS Annual Survey File Unconditional Response Rates for Continuously-Enrolled Beneficiaries

Panel	Released	Continuously -Enrolled Complete	Continuously -Enrolled Eligible	Continuously -Enrolled Ineligible	Unconditional Response Rate for Continuously-Enrolled Beneficiaries
2012	7,400	2,153	6,484	916	33.2%
2013	7,400	2,457	6,608	792	37.2%
2014	11,398	4,188	10,489	909	39.9%
2015	8,621	4,134	8,179	442	50.5%
Total	34,819	12,932	31,760	3,059	40.7%

Exhibits 9.1.5 and 9.1.6 display the 2015 Survey File conditional response rates by panel for ever-enrolled and continuously-enrolled beneficiaries.

Exhibit 9.1.5: 2015 MCBS Annual Survey File Conditional Response Rates for Ever Enrolled Beneficiaries

Panel	Ever-Enrolled Complete	Ever-Enrolled Eligible	Subset of Ever-Enrolled Eligibles That Were Not Released	Conditional Response Rate for Ever-Enrolled Beneficiaries
2012	2,380	7,298	4,140	75.4%
2013	2,716	7,300	3,389	69.4%
2014	4,626	11,255	4,923	73.1%
2015	4,349	8,179	0	53.2%
Total	14,071	34,032	12,452	65.2%

Exhibit 9.1.6: 2015 MCBS Annual Survey File Conditional Response Rates for Continuously-Enrolled Beneficiaries

Panel	Continuously-Enrolled Complete	Continuously-Enrolled Eligible	Subset of Continuously-Enrolled Eligibles That Were Not Released	Conditional Response Rate for Continuously-Enrolled Beneficiaries
2012	2,153	6,484	3,451	71.0%
2013	2,457	6,608	2,874	65.8%
2014	4,188	10,489	4,421	69.0%
2015	4,134	8,179	0	50.5%
Total	12,932	31,760	10,746	61.5%

9.2 Nonresponse Bias Analysis

Survey nonresponse occurs when data are not collected for an eligible sampled individual. Unit nonresponse occurs when no data are collected for a sampled individual; item nonresponse occurs when data for a particular questionnaire item or items are not collected from a sampled respondent. If respondents differ from non-respondents in meaningful ways, then nonresponse bias may occur.

This section discusses several analyses[44] that were conducted to evaluate whether and how much nonresponse bias is evident in the MCBS. It is presented in four parts. First, MCBS attrition rates across several rounds of data collection are presented to give a sense of the level of nonresponse to the survey. Second, respondents are compared to non-respondents on a variety of measures available for all sampled individuals. Fall 2015 respondents are compared to non-respondents based on frame data available for all beneficiaries sampled into the MCBS. A logistic regression model is also developed and analyzed as an additional means of evaluating nonresponse bias based on frame characteristics. Third, comparisons of Fall 2015 respondents and non-respondents using various claims payment measures are presented. Finally, the differences between respondents and nonrespondents with respect to various chronic conditions indicators are studied, all of which are available for all Medicare beneficiaries. The section concludes with a brief summary of findings.

MCBS Attrition Rates

Exhibit 9.2.1 displays MCBS panel attrition by round for beneficiaries in the 2015 Survey File. Note that in Summer 2016, some cases were intentionally not fielded and instead were included in an early case release for the Fall 2016 round. This provided a way to better balance case releases in future rounds and start the Fall round earlier (which is always important as the Fall round includes

---

[44] Analyses presented in this section are unweighted.

the new Incoming Panel). To account for this, two additional rows of data—deferred and released for fieldwork—are provided for each panel. The 'deferred' row displays the number of cases intentionally not fielded in the round (zero for all rounds except Summer 2016), and the second row displays the number of cases released for fieldwork in the round (equal to the panel total in all rounds except Summer 2016). The calculation of the attrition rate is based on the number of cases released for fieldwork rather that the total for the round.

Generally, a panel is expected to experience the most attrition in its first round in data collection, and attrition is expected to decrease steadily for each panel over time. This appears to be the trend for MCBS as well, as evidenced by the data shown in Exhibit 9.2.2. This exhibit presents attrition rates in the Fall of 2015 and into the following four[45] rounds. The incoming (2015) panel had a nearly 50 percent attrition rate in its first round in the field (Fall 2015), but that attrition tapered off quickly to below 20 percent in each of the next three rounds. Other panels had much lower attrition rates in Fall 2015, with attrition increasing slightly over the subsequent rounds before dropping again to below the Fall 2015 rate by Winter 2017. The 2012 Panel, which was retired in Summer 2016, actually saw its attrition rates increase slightly over its last few rounds in the field.

Exhibit 9.2.1: Unweighted Attrition Rates by Panel, Fall 2015 to Winter 2017: 2015 Survey File

	Fall 2015	Winter 2016	Summer 2016	Fall 2016	Winter 2017
2015 Panel					
Deferred	0	0	571	0	0
Released for Fieldwork	8,621	4,349	3,027	3,084	2,488
Non-Respondents	4,272	751	514	596	256
Respondents	4,349	3,598	2,513	2,488	2,232
Total	8,621	4,349	3,598	3,084	2,488
2015 Panel Attrition Rate[1]	49.6%	17.3%	17.0%	19.3%	10.3%
2014 Panel					
Deferred	0	0	1,043	0	0
Released for Fieldwork	5,276	4,626	2,818	3,292	2,753
Non-Respondents	650	765	569	539	241
Respondents	4,626	3,861	2,249	2,753	2,512
Total	5,276	4,626	3,861	3,292	2,753
2014 Panel Attrition Rate[1]	12.3%	16.5%	20.2%	16.4%	8.8%
2013 Panel					
Deferred	0	0	690	0	0
Released for Fieldwork	3,003	2,716	1,659	2,031	1,790
Non-Respondents	287	367	318	241	109

---

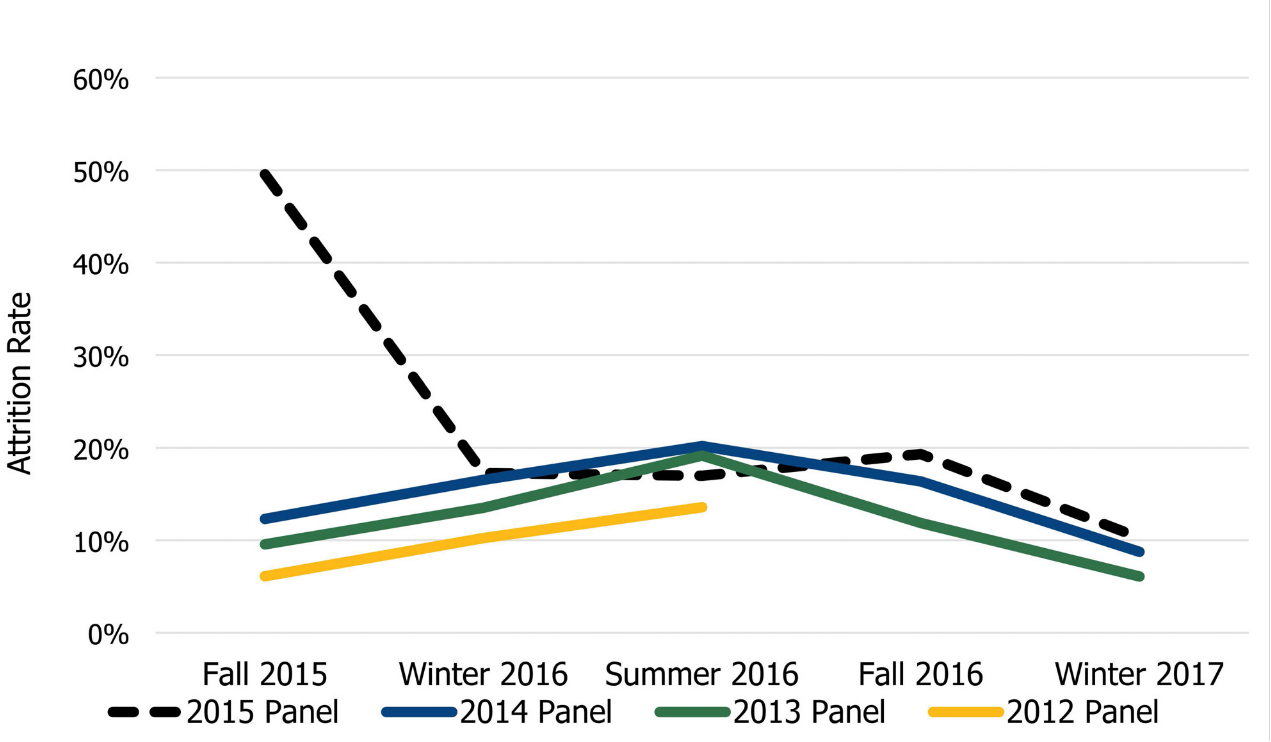
[45] Due to the timing of this report, attrition rates are available for rounds beyond 2015. Including them here helps to illustrate the trends a panel experiences over time.

	Fall 2015	Winter 2016	Summer 2016	Fall 2016	Winter 2017
Respondents	2,716	2,349	1,341	1,790	1,681
Total	3,003	2,716	2,349	2,031	1,790
2013 Panel Attrition Rate[1]	9.6%	13.5%	19.2%	11.9%	6.1%
2012 Panel					
Deferred	0	0	0		
Released for Fieldwork	2,535	2,380	2,136		
Non-Respondents	155	244	290		
Respondents	2,380	2,136	1,846		
Total	2,535	2,380	2,136		
2012 Panel Attrition Rate[1]	6.1%	10.3%	13.6%		
Continuing Sample, 2012-2014 Panels					
Deferred	0	0	1,733	0	0
Released for Fieldwork	10,814	9,722	6,613	5,323	4,543
Non-Respondents	1,092	1,376	1,177	780	350
Respondents	9,722	8,346	5,436	4,543	4,193
Total	10,814	9,722	8,346	5,323	4,543
Attrition Rate[1]	10.1%	14.2%	17.8%	14.7%	7.7%

[1] Calculated as Non-Respondents/Released for Fieldwork.

Exhibit 9.2.2 provides a visual display of these rates over time.

Exhibit 9.2.2: Unweighted Attrition Rates by Panel, Fall 2015 to Winter 2017: 2015 Survey File



Comparison of Respondents to Non-Respondents: Frame Characteristics

As displayed above, the response rate for the 2015 Panel in Fall 2015 was 50.4 percent, and the continuing panels' Fall 2015 response rates averaged around 90 percent. Both respondents and non-respondents from each panel are included in their respective frame files, from which the samples were selected. The frame files contain a variety of demographic, location, and other background information for all beneficiaries eligible to be sampled as part of that panel. Therefore, it is possible to compare respondents and non-respondents based on these frame characteristics. A comparison can help to detect noticeable differences between these two groups and perhaps identify areas of potential bias resulting from nonresponse.

Exhibit 9.2.3 displays comparisons of 2015 Panel Fall 2015 respondents to non-respondents using several of the frame characteristics, including sex, age, race/ethnicity, current-year enrollee (i.e., whether the beneficiary became eligible and enrolled during 2015), Census division, Health and Human Services (HHS) Region[46], and ACO status (i.e., whether the beneficiary was enrolled in an Accountable Care Organization; ACO members were previously oversampled in the MCBS). (Note that the comparisons that follow include only community-dwelling beneficiaries; facility-dwelling

---

[46] Regions defined for the purposes of program and outreach coordination for the U.S. Department of Health and Human Services.

beneficiaries are excluded by design. Thus, the table counts will be smaller than for the attrition rates table presented earlier in this section.) To test differences between the two populations, the Rao-Scott chi-square test was used. This test adjusts the Pearson Chi-Square statistics, using a second-order design correction, by dividing it twice by the generalized design effect factor (GDEFF). The second-order correction adjusts not only the mean of the chi-square distribution but also the variance.

Statistically significant differences between respondents and non-respondents were detected for sex, age, race/ethnicity, Census division, and HHS region. While non-respondents appear more likely to be female and older, and slightly less likely to be non-Hispanic black, the differences are not large. It is always advisable to take caution when interpreting significant differences, as large sample sizes such as those in the MCBS can result in statistically significant differences being found even when little practical difference is observed. Thus, it is important to observe the actual differences in percentages between respondents and non-respondents within variable categories to identify practical differences between the two groups.

Exhibit 9.2.3: 2015 Panel Respondents vs. Non-respondents in Fall 2015, by Frame Characteristics

	Fall 2015 Non-Respondents #	Fall 2015 Respondents #	Fall 2015 Non-Respondents %	Fall 2015 Respondents %
Sex:**				
Male	1,820	1,976	42.6%	45.4%
Female	2,444	2,373	57.2%	54.6%
Missing	8	0	0.2%	0.0%
Age:**				
Under 45	418	420	9.8%	9.7%
45-64	212	316	5.0%	7.3%
65-69	634	615	14.8%	14.1%
70-74	685	660	16.0%	15.2%
75-79	691	784	16.2%	18.0%
80-84	748	754	17.5%	17.3%
85 and over	884	800	20.7%	18.4%
Race/Ethnicity:**				
Hispanic	391	409	9.2%	9.4%
Non-Hispanic White	2,977	3,030	69.7%	69.7%
Non-Hispanic Black	344	423	8.1%	9.7%
All Other	180	108	4.2%	2.5%
Missing/Unknown	380	379	8.9%	8.7%

	Fall 2015 Non-Respondents #	Fall 2015 Respondents #	Fall 2015 Non-Respondents %	Fall 2015 Respondents %
Current-Year Enrollee:				
Not Current-Year Enrollee	3,911	3,985	91.5%	91.6%
Current-Year Enrollee	361	364	8.5%	8.4%
Census Division:**				
Northeast	795	785	18.6%	18.1%
Midwest	877	1,019	20.5%	23.4%
South	1,611	1,741	37.7%	40.0%
West	989	804	23.2%	18.5%
HHS Census Region:**				
1: CT, MA, ME, NH, RI, VT	162	127	3.8%	2.9%
2: NJ, NY, PR	630	655	14.7%	15.1%
3: DC, DE, MD, PA, VA, WV	645	705	15.1%	16.2%
4: AL, FL, GA, KY, MS, NC, SC, TN	236	316	5.5%	7.3%
5: IL, IN, MI, MN, OH, WI	854	898	20.0%	20.6%
6: AR, LA, NM, OK, TX	300	347	7.0%	8.0%
7: IA, KS, MO, NE	416	425	9.7%	9.8%
8: CO, MT, ND, SD, UT, WY	348	354	8.1%	8.1%
9: AZ, CA, HI, NV	643	450	15.1%	10.3%
10: AK, ID, OR, WA	38	72	0.9%	1.7%
ACO Status:				
Not ACO	3,634	3,666	85.1%	84.3%
ACO	638	683	14.9%	15.7%

Source: 2015 Survey File.

**: Statistically significant at P<.01

Exhibit 9.2.4 displays comparisons of the combined 2012 through 2014 (continuing) Panel respondents to non-respondents based on the same frame characteristics[47]. For the continuing panels, most of the distributions across the various frame variables were similar for respondents and non-respondents. The only statistically significant differences detected between respondents and non-respondents were within the age categories; the non-respondents tend to skew younger than the respondents.

---

[47] The only exception to this is race/ethnicity. Prior to 2014, the only race/ethnicity data available for beneficiaries was a black/not black indicator. Beginning in 2014, a more detailed racial/ethnic variable was delivered as part of the beneficiary frame file. Thus, the race/ethnicity categories for Exhibit 9.2.4 are more limited than for Exhibit 9.2.3.

Exhibit 9.2.4: 2012-2014 Panel Respondents vs. Non-respondents in Fall 2015, by Frame Characteristics

	Fall 2015 Non-Respondents #	Fall 2015 Respondents #	Fall 2015 Non-Respondents %	Fall 2015 Respondents %
Sex:				
Male	479	4,308	43.9%	44.3%
Female	613	5,414	56.1%	55.7%
Age*:				
Under 45	134	831	12.3%	8.5%
45-64	91	815	8.3%	8.4%
65-69	251	2,153	23.0%	22.1%
70-74	149	1,413	13.6%	14.5%
75-79	170	1,581	15.6%	16.3%
80-84	163	1,533	14.9%	15.8%
85 and over	134	1,396	12.3%	14.4%
Race[a]:				
Black	118	1,035	10.8%	10.6%
Not Black	974	8,687	89.2%	89.4%
Census Division:				
Northeast	194	1,859	17.8%	19.1%
Midwest	312	2,467	28.6%	25.4%
South	377	3,657	34.5%	37.6%
West	209	1,739	19.1%	17.9%
HHS Census Region:				
1 CT, MA, ME, NH, RI, VT	48	339	4.4%	3.5%
2 NJ, NY, PR	139	1,420	12.7%	14.6%
3 DC, DE, MD, PA, VA, WV	188	1,788	17.2%	18.4%
4 AL, FL, GA, KY, MS, NC, SC, TN	124	685	11.4%	7.0%
5 IL, IN, MI, MN, OH, WI	185	1,944	16.9%	20.0%
6 AR, LA, NM, OK, TX	82	710	7.5%	7.3%
7 IA, KS, MO, NE	104	922	9.5%	9.5%
8 CO, MT, ND, SD, UT, WY	84	756	7.7%	7.8%
9 AZ, CA, HI, NV	125	980	11.4%	10.1%
10 AK, ID, OR, WA	13	178	1.2%	1.8%
ACO:				
Not ACO	709	5,604	75.7%	76.3%
ACO	228	1,738	24.3%	23.7%

Source: 2015 Survey File.

[a] Prior to 2014, the only race/ethnicity information available for beneficiaries is a Black/Not Black indicator.

*: Statistically significant at P<.05

Next, hard-to-contact (HR) respondents, defined here as respondents with at least nine contact attempts in the field, were separated from other respondents. Approximately 10 percent of the respondents fall into the hard-to-contact category. Then, the 2015 Panel Fall 2015 non-respondents (NR) are compared to both hard-to-contact respondents and other respondents across the frame characteristics described above. Exhibit 9.2.5 presents the number and proportion of non-respondents, hard-to-contact respondents, and other respondents with each frame characteristic. Non-respondents appear to be more likely to be female (57.8 percent) compared to hard-to-contact (53.1 percent) and other (53.8 percent) respondents. There are also some noticeable differences across age and race/ethnicity, especially between hard-to-contact respondents and the other two groups. For example, the hard-to-contact respondents tend to skew younger than the non-respondents and other respondents and are more likely to be Hispanic and non-Hispanic black (and less likely to be non-Hispanic white) than the other two groups. Some statistically significant differences were detected between non-respondents and hard-to-contact respondents, and between non-respondents and other respondents. These are indicated in the top row of the percent columns for each characteristic. Again, because the MCBS sample size is very large, these significance levels do not necessarily indicate meaningful or practical differences.

Exhibit 9.2.5: 2015 Panel Non-respondents (NR) vs. Hard-to-Contact Respondents (HR) and Other Respondents (OR), by Frame Characteristics: Fall 2015

Frame Characteristic	NR #	HR #	OR #	NR %	HR %	OR %
Sex:						**
Male	1,594	176	1,721	42.2%	46.9%	46.2%
Female	2,180	199	2,001	57.8%	53.1%	53.8%
Age:					**	**
Under 45	375	55	345	9.9%	14.7%	9.3%
45-64	197	29	275	5.2%	7.7%	7.4%
65-69	615	72	538	16.3%	19.2%	14.5%
70-74	661	62	583	17.5%	16.5%	15.7%
75-79	637	54	707	16.9%	14.4%	19.0%
80-84	664	47	672	17.6%	12.5%	18.1%
85 and over	625	56	602	16.6%	14.9%	16.2%
Race/Ethnicity:					**	**
Hispanic	343	43	348	9.1%	11.5%	9.3%
Non-Hispanic White	2,586	227	2,600	68.5%	60.5%	69.9%
Non-Hispanic Black	300	53	345	7.9%	14.1%	9.3%
All Other	174	14	89	4.6%	3.7%	2.4%
Missing/Unknown	371	38	340	9.8%	10.1%	9.1%
Current-Year Enrollee:						
Not Current-Year Enrollee	3,418	339	3,495	90.6%	90.4%	91.4%
Current-Year Enrollee	356	36	327	9.4%	9.6%	8.6%

Frame Characteristic	NR #	HR #	OR #	NR %	HR %	OR %
Census Division:						**
Northeast	687	84	645	18.2%	22.4%	17.3%
Midwest	772	69	875	20.5%	18.4%	23.5%
South	1,418	140	1,525	37.6%	37.3%	41.0%
West	897	82	677	23.8%	21.9%	18.2%
HHS Census Region:						**
1 CT, MA, ME, NH, RI, VT	140	17	101	3.7%	4.5%	2.7%
2 NJ, NY, PR	543	66	542	14.4%	17.6%	14.6%
3 DC, DE, MD, PA, VA, WV	565	44	613	15.0%	11.7%	16.5%
4 AL, FL, GA, KY, MS, NC, SC, TN	209	26	263	5.5%	6.9%	7.1%
5 IL, IN, MI, MN, OH, WI	757	70	784	20.1%	18.7%	21.1%
6 AR, LA, NM, OK, TX	258	36	301	6.8%	9.6%	8.1%
7 IA, KS, MO, NE	370	35	370	9.8%	9.3%	9.9%
8 CO, MT, ND, SD, UT, WY	310	37	303	8.2%	9.9%	8.1%
9 AZ, CA, HI, NV	590	44	375	15.6%	11.7%	10.1%
10 AK, ID, OR, WA	32	-	70	0.8%	0.0%	1.9%
ACO:						
Not ACO	3,216	314	3,150	85.2%	83.7%	84.6%
ACO	558	61	572	14.8%	16.3%	15.4%

[a] Beneficiaries living in the community only; beneficiaries living in facilities excluded by design.
* Significant at P<0.05
** Significant at P<0.01

As an additional means of evaluating nonresponse based on frame data, multivariate analyses were used to identify the characteristics of beneficiaries least likely to respond to an interview. Cases were coded as either responding to or not responding to the Fall 2015 interview, and logistic regression modeling was used to identify which case characteristics significantly relate to unit nonresponse.

For the 2015 panel, frame attributes were used as covariates to build a logistic regression model of Fall 2015 non-respondents. The dependent variable is an indicator identifying whether the beneficiary is a non-respondent. The independent variables include sex, race/ethnicity, age group, a current-year enrollee flag, Census division, and a hard-to-contact indicator for all beneficiaries (coded as 1 for any beneficiary, respondent or non-respondent, requiring at least 9 contact attempts in the field, and 0 otherwise). We used the stepwise option for model selection; only the current-year enrollee flag was dropped from the model. Since multiple tests are performed, in order to prevent falsely significant results we used the Bonferroni adjustment to adjust p-values for multiplicity and test significant differences.

Exhibit 9.2.6 displays the results of the logistic regression analysis. Holding all other covariates at a fixed value, the odds of being a non-respondent is about 14 percent lower for males (0.86 odds ratio) than for females. Beneficiaries of other races[48] are almost twice as likely to be non-respondents as Hispanic beneficiaries. Furthermore, younger beneficiaries (age 45-64) are 31 percent less likely to be non-respondents than beneficiaries aged 85 and older, and beneficiaries who live in the South are 32 percent more likely to be non-respondents than those in the West. Finally, hard-to-contact beneficiaries are about 60 percent more likely to be non-respondents than other beneficiaries.

Exhibit 9.2.6: Logistic Regression Model of 2015 Panel Non-respondents, Fall 2015

Effect	Estimates	Standard Error	Odds Ratio Estimates	Adjusted P Value
SEX Male vs Female	-0.16	0.05	0.86	0.0009
RACE Non-Hispanic White vs Hispanic	0.14	0.08	1.15	1.0000
RACE Non-Hispanic Black vs Hispanic	-0.12	0.11	0.89	0.5790
RACE Other vs Hispanic	0.68	0.15	1.97	<.0001
AGE Group <45 vs 85+	-0.05	0.10	0.95	1.0000
AGE Group 45-64 vs 85+	-0.38	0.11	0.69	0.0140
AGE Group 65-69 vs 85+	-0.02	0.09	0.98	1.0000
AGE Group 70-74 vs 85+	0.04	0.08	1.04	1.0000
AGE Group 75-79 vs 85+	-0.14	0.08	0.87	1.0000
AGE Group 80-84 vs 85+	-0.03	0.08	0.97	1.0000
CENSUS Division North East vs West	0.03	0.07	1.03	1.0000
CENSUS Division Mid- West vs West	-0.06	0.06	0.94	1.0000
CENSUS Division South vs West	0.28	0.06	1.32	<.0001
Hard to Contact No vs Yes	-0.89	0.06	0.41	<.0001

Comparison of Respondents to Non-Respondents: Medicare Claims Payment Measures

For the next set of analyses, non-respondents, hard-to-contact respondents, and other respondents were compared using 2015 claims data to identify any differences in claims payment amounts among these groups. In order to conduct the composite test for no differences among the three response categories, we used a generalized linear model (GLM), with the claims payment amount as the dependent variable and a three-level response indicator (hard-to-contact respondent, other respondent, or non-respondent) as the independent variable. Contrast statements were used in a one-way ANOVA to test the hypothesis that no differences exist among these groups. Exhibit 9.2.7 shows a comparison of 2015 Panel non-respondents, hard-to-contact respondents, and other respondents in Fall 2015 across seven claims payment amount categories. Mean payment amounts for each group are presented in the first three columns, and significant

---

[48] Defined as beneficiaries not coded as Hispanic, Non-Hispanic White, or Non-Hispanic Black.

differences for the three contrast comparisons are indicated in the subsequent columns. No significant differences were found among any of the groups.

Exhibit 9.2.7: 2015 Claims Payment Measures for 2015 Panel[a] Non-Respondents (NR), Hard-to-Contact Respondents (HR), and Other Respondents (OR): Fall 2015

Claims Payment Measures	Mean of NR $	Mean of OR $	Mean of HR $
Claims payment amount: Carrier	1,924.52	2,025.22	2,137.68
Claims payment amount: Durable medical equipment	132.59	146.53	114.51
Claims payment amount: Home health agency	320.13	374.60	466.56
Claims payment amount: Hospice	168.36	194.42	12.52
Claims payment amount: Inpatient	1,810.45	1,919.52	2,472.34
Claims payment amount: Outpatient	938.93	1,155.89	1,060.75
Claims payment amount: Skilled nursing facility	419.83	374.20	420.36
All claims: Total payment amount	5,714.72	6,190.34	6,684.72

[a] Beneficiaries living in the community only; beneficiaries living in facilities excluded by design.

Comparison of Respondents to Non-Respondents: Chronic Condition Attributes

A final source of data used in this nonresponse bias analysis is obtained from the Chronic Conditions Warehouse. These data are available on an annual basis for all Medicare beneficiaries and identify whether a beneficiary met the claims and/or coverage criteria to be classified as having a particular chronic condition. These conditions include chronic kidney disease, diabetes, depression, stroke, breast cancer, anemia, asthma, and benign prostatic hyperplasia.

For this analysis, Fall 2015 respondents were compared to non-respondents across 2015 year end Chronic Condition attributes. Exhibit 9.2.8 displays the percentage of 2015 Panel and continuing panel respondents and non-respondents classified as having each particular chronic condition, based on meeting the claims and/or coverage criteria for each condition. Again, the Rao-Scott chi-square test was used to test the significance of differences between respondents and non-respondents. No significant differences in chronic conditions attributes were found between respondents and non-respondents in either the 2015 or combined 2012-2014 Panels.

Exhibit 9.2.8: Fall 2015 Respondents vs. Non-Respondents, by Round, Across Chronic Condition Attributes

Chronic Condition	2015 Panel[a]		2012-2014 Panels[a]	
	Non-Respondent	Respondent	Non-Respondent	Respondent
Chronic Kidney Disease	58.1	58.2	62.3	63.2
Diabetes	59.7	59.9	63.5	64.6
Depression	60.9	60.8	63.5	64.3
Stroke/Transient Ischemic Attack	59.9	59.8	62.7	63.4
Breast Cancer	59.7	59.7	62.6	63.3
Anemia	61.2	61.1	63.4	64.7
Asthma	60.2	60.0	62.8	63.7
Benign Prostatic Hyperplasia	59.9	59.9	63.0	63.5

[a] Beneficiaries living in the community only; beneficiaries living in facilities excluded by design.

Summary and Implications

Attrition rates in the MCBS follow patterns typical to longitudinal studies, with the highest attrition occurring at the first time in sample and attrition rates decreasing over subsequent rounds. The attrition rate for the 2015 Panel was just under 50 percent in Fall 2015 and decreased to approximately 10 percent by Winter 2017.

Fall 2015 respondents and non-respondents were compared on various measures, including frame characteristics, Medicare claims payments, and chronic conditions, in order to identify areas of potential bias. The only statistically significant differences were found among frame characteristics. For the 2015 Panel, non-respondents appear more likely to be female and older, and slightly less likely to be non-Hispanic black. Among the continuing panels, however, non-respondents tend to skew younger. Furthermore, the weighting procedure includes a raking step that accounts for all of the frame characteristics for which differences were found (see Section 8.3; raking to control totals is performed using several frame variables, including age group, gender, race and census region, among others; an additional nonresponse bias adjustment is also made). Thus, the small potential bias identified via these analyses is expected to be minimized by the weighting procedures.

10. USING MCBS DATA FILES

10.1 MCBS Data User's Guide

The MCBS Data User's Guide offers a publicly available, easily searchable resource for data users. Beginning with 2015 MCBS data, it is updated for each new data year to ensure that users have current documentation on the survey design, questionnaires, and estimation as well as detailed notes on the structure and contents of the MCBS data releases.

The Data User's Guide features three stand-alone chapters. The General Data User's Guide documents the key features of the study and data products. The Survey File and Cost Supplement Guides provide technical information on each file including the derivation of variables and any significant changes in the variables and/or file structure. The Data User's Guides can be located on CMS' MCBS website at https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS/Codebooks.html.

10.2 MCBS Public Use Data File

Beginning with data collected in the 2013 MCBS, a public use file (PUF) and accompanying documentation are available free for download under the MCBS PUF link at https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/MCBS-Public-Use-File/index.html. The MCBS PUF is an easy to use data file with select data items that allow researchers to conduct analysis on health disparities, access to and satisfaction with healthcare, and medical conditions for community dwelling Medicare beneficiaries. The MCBS PUF is not intended to replace the more detailed limited data set (LDS) files, rather it provides a publically available alternative for those researchers interested in the health, health care use, access to and satisfaction with Medicare of beneficiaries, while providing the very highest degree of protection to the Medicare beneficiaries' protected health information, meeting all necessary de-identification of the data and mitigating disclosure risk.

10.3 MCBS Limited Data Sets

There are two MCBS Limited Data Sets (LDS) available to data users. In order to access these data files, data users must submit a Data Use Agreement (DUA) and complete an LDS Worksheet, which provides CMS with information about the research project, the particular files needed, and payment information for administrative fees associated with the data request. Note that new data users and repeat data users complete distinct forms. Data users should visit CMS' LDS website at https://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/Data-Disclosures-Data-Agreements/DUA_-_NewLDS.html for a full description of available LDS files, further details about the request process, and downloadable forms.

11. GLOSSARY

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: An individual selected from MCBS' sample about whom the MCBS collects information. Beneficiary may also refer to a person receiving Medicare services who may or not be participating in the MCBS.

Claim-only event: A claim-only event is a medical service or event known only through the presence of a Medicare claim. The event did not originate from an event or service reported by a respondent during an interview.

Community component: Survey of beneficiaries residing in the community at the time of the interview (i.e., not in a long-term care facility such as a nursing home). Beneficiaries answered health status and functioning questions themselves, unless they were unable to do so.

Continuing interview: The questionnaire administered to repeat respondents as they progress through the study (interviews #2-12).

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 other than Puerto Rico.

Core modules: These sections of the MCBS Questionnaire are of critical purpose and policy relevancy to the MCBS, regardless of season of administration.

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

Ever enrolled: A Medicare beneficiary who was enrolled at any time during the calendar year including those 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 other than Puerto Rico.

Exit interview: Conducted in the summer round, this interview completes the respondent's participation in the MCBS (interview #12). The exit interview is a special case of the Continuing interview.

Facility component: Survey of beneficiaries residing in facilities, such as long-term care nursing homes or other institutions, at the time of the interview. Facility interviewers do not conduct the Facility component with the respondent, but with a staff member located at the facility. This is a key difference between the Community and Facility components.

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.

Incoming Panel Sample (formerly known as Supplemental Panel): A scientifically selected group of sampled beneficiaries that enter the MCBS in the fall of a data collection year. One panel is retired during each summer 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.

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 (HCPPs).

Medicare beneficiary (aka, beneficiary): An individual who meets at least one of three criteria (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).

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.html.

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 medicines: The basic unit measuring use of prescription medicines is a single purchase of a single drug in a single container. Prescription drug use is collected only for beneficiaries living in the community or in a facility, and does not include prescription medicines administered during an inpatient hospital stay.

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.

Race/ethnicity: Responses to race and ethnicity questions were recorded as interpreted by the respondent. Respondents who reported they were white and not of Hispanic origin were coded as white non-Hispanic; those who reported they were black/African-American and not of Hispanic origin were coded as black non-Hispanic; persons who reported they were Hispanic, regardless of their race, were coded as Hispanic; persons who reported they were American Indian, an Asian or Pacific Islander, or other race and not of Hispanic origin were coded as other race/ethnicity. Hispanic includes persons of Mexican, Puerto Rican, Cuban, Central or South American, or other Spanish culture or origin, regardless of race. Respondents with more than one racial background were captured in a separate category and collapsed into the "other" category.

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

Residence status: Full-year community residents are Medicare beneficiaries who lived solely in household units during the data collection year and who received community interviews only. Full-year facility residents are Medicare beneficiaries who lived solely in a long-term care facility during the data collection year and who received Facility interviews only. Part-year community/part-year facility residents are Medicare beneficiaries who lived part of the year in the community and part of the year in a long-term care facility, and who received both Community and Facility interviews. Skilled nursing facility users are Medicare beneficiaries who lived in either the community or a facility, and who used skilled nursing facility services during the data collection year.

Respondent: The person who answers questions about the beneficiary for the MCBS; this person can be the beneficiary themselves, 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).

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 claim, 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.

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.

Copyright Information and Guidelines for Data Use

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.

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, 2015.