

Medicare Current Beneficiary Survey
CY 2006 Access to Care

Public Use File Documentation

Introduction

The accompanying data file is the sixteenth in a series of annual data releases relating to Medicare beneficiaries' access to care. The file consists of selected interview data from the ongoing Medicare Current Beneficiary Survey (MCBS), which were collected during Round 46 (September through December of 2006) or earlier rounds for some variables for individuals in the continuing sample. These data are augmented with Medicare claims and administrative data for calendar year (CY) 2006.

Purposes of the Survey

The MCBS is a continuous, multi-purpose panel survey of a representative sample of the Medicare population, including both aged and disabled enrollees. The study is sponsored by the Centers for Medicare and Medicaid Services (CMS). CMS' primary mission is administering the Medicare program and assisting the States in administering the Medicaid program. The MCBS is designed to support the agency's functions by providing data necessary for policymakers to develop needed program enhancements.

U.S. health expenditures grew 6.7 percent in 2006 to \$2.1 trillion, up slightly from 6.5 percent growth in 2005. On a per capita basis, health spending increased by \$377 to \$7,026. Health spending accounted for 16 percent of Gross Domestic Product in 2006, remaining relatively stable as a share since 2003.

After six consecutive years of slowing growth retail prescription drug spending growth accelerated in 2006, at the same time, implementation of the Medicare Part D benefit caused a major shift in the sources of funds used to pay for drugs. In contrast, in 2006, there was a broadly based slowdown across most major health care services, including hospital, physician and clinical, dental, other professional, home health and nursing home care services. Private spending slowed partially offsetting the slight acceleration in growth in public funding in 2006.

These trends, along with CMS' concerns about the quality of care beneficiaries receive; support the need for the collection of a wide variety of health related data. The MCBS is designed to aid in CMS' administration, monitoring and evaluation of the Medicare and Medicaid programs. MCBS data enables CMS to do the following: monitor the impact, especially financial, of changes in the Medicare program on the beneficiary population and on the trust funds; develop reliable and current information on the use and cost of services not covered by Medicare (such as long-term care); develop reliable and current information on the sources of payment for costs of covered services not reimbursed by Medicare and of noncovered services; and analyze factors which are not available from claims or other administrative records but are thought to affect use and mix of services.

The MCBS primarily focuses on economic and beneficiary issues; in particular, health care use, expenditures and factors that affect use of care and the beneficiary's ability to pay. As a part of this focus the MCBS collects a variety of information about demographic characteristics, health status and functioning, access to care, insurance coverage, financial resources and potential family support. The longitudinal design of the MCBS allows analysis of the effects of changes in these factors on patterns of use over time.

The Design of the MCBS

CMS conducts the MCBS through a data collection contractor.

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

Fieldwork for Round 1 began in September 1991 and was completed in December 1991. Subsequent rounds, involving the re-interviewing of the same sample persons or appropriate proxy respondents, begin every four months. Interviews are conducted regardless of whether the sample person resides at home or in a long-term care facility, using the questionnaire version (discussed later) appropriate to the setting.

Repeated Interviews. The MCBS is a longitudinal panel survey. Sample persons are interviewed three times a year for a maximum of four years to form a continuous profile of each individual's personal health care experience. The MCBS is thus uniquely capable of tracing changes in coverage and other personal circumstances and observing processes that occur over time, such as people leaving their homes and taking up residence in long-term care facilities or spending down their assets for medical care until they become eligible for Medicaid.

Sample. Respondents for the MCBS were sampled from the Medicare enrollment file to be representative of the Medicare population as a whole and by the following age groups: under 45, 45 to 64, 65 to 69, 70 to 74, 75 to 79, 80 to 84, and 85 and over. Because of interest in their special health care needs, the oldest old (85 and over) and the disabled (64 and under) were oversampled to permit more detailed analysis of these subpopulations. The MCBS sample was designed to yield about 16,000 completed cases annually in the Access to Care data file.

The sample was drawn from 107 primary sampling units (PSUs) or major geographic areas chosen to represent the nation, including the District of Columbia and Puerto Rico, with a second stage of 1,163 geographic clusters defined by ZIP Code initially drawn within those PSUs. The second-stage sample was expanded each subsequent year, increasing to 1,523 in Round 28. For Round 31, the PSU sample was updated and reselected.

Within the new sample of 107 PSUs, 1,209 geographic clusters were initially selected in Round 31. With the addition of new ZIP code clusters in subsequent years, the number of geographic clusters increased to 1,244 in Round 46. The MCBS sample is annually supplemented during the September through December interview periods (that is, Rounds 37, 40, 43, and 46) to account for attrition (deaths, disenrollments, refusals, etc.) and newly enrolled persons. These annual supplements are referred to as panels.

The beneficiaries included in the 2006 Access to Care File consist of a random cross-section of all beneficiaries who were enrolled in one or both parts of the Medicare program as of January 1, 2006 and were alive and enrolled at the time of interview during the 2006 fall round (September - December). These beneficiaries include those in four separate MCBS panels (the 2003, 2004, 2005 and 2006 panels) and were drawn through the use of a fairly complex selection algorithm.

The following bullets describe panel composition for each of the four panels participating in this data file.

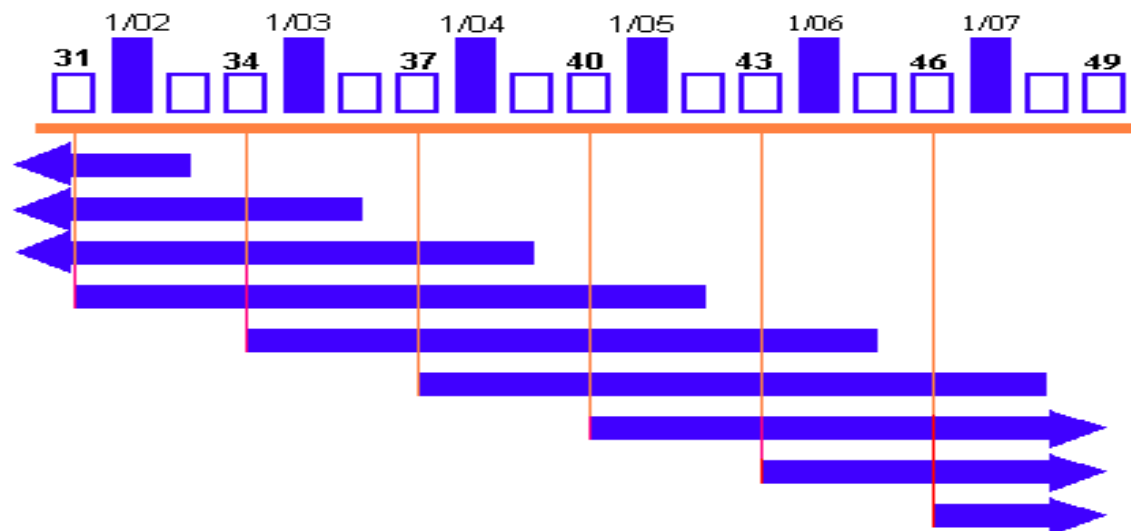
- In Round 37 (September-December 2003) the tenth rotating panel was selected, consisting of 5,930 beneficiaries. This panel was the third to be selected from the new (updated) sample of PSUs.
- In Round 40 (September-December 2004) the eleventh rotating panel was selected, consisting of 5,978 Beneficiaries. This panel was the fourth to be selected from the new (updated) sample of PSUs.
- In Round 43 (September-December 2005) the twelfth rotating panel was selected, consisting of 6,565 Beneficiaries. This panel was the fifth to be selected from the new (updated) sample of PSUs.
- In Round 46 (September-December 2006) the thirteenth rotating panel was selected, consisting of 6,675 Beneficiaries. This panel was the sixth to be selected from the new (updated) sample of PSUs. In addition to the 2006 panel, a special supplement of 1,490 newly enrolled Medicare beneficiaries was fielded in Round 46.

A rotating panel will be followed for 12 interviews. There are four panels active at any one time, and each panel has approximately 4,000 active sample persons. New panels that are introduced each year in the fall round will replace the oldest panel that will subsequently be retired in the following summer.

Because the access to care data file includes four panels, the number of interviews conducted in the September—December (Fall) round is approximately 16,000. Figure I.1, while not drawn to scale, gives a visual display of the overlap that occurs during the simultaneous fielding of four panels in the fall round.

The retiring panel (i.e., the 2002 panel consisting of about 4,000 individuals) has abbreviated questionnaires administered beginning in the January to April round of 2006 and, if necessary, the following May to August (summer) round to complete the collection of medical events occurring in the previous calendar year. These sample persons were then rotated out of the study. These individuals participated a maximum of four years (that is, a baseline interview, three complete years of utilization and expenditure data, and up to two interviews to "close out" events due to late arriving paperwork).

During each Fall round, under the rotating panel design, a new panel will be introduced and during each following Summer round a panel will be retired. Thus, for example, the new panel that was introduced in Round 46 will replace approximately 4,000 of the ongoing sample (i.e., the 2003 panel) by Round 48. This rotating panel sample design allows for both the eventual termination of participation in the study for individuals and for the completion of about 16,000 interviews in each Fall round for the ongoing study population.

Figure I.1: Rotating Panel Overlap

As noted earlier, Figure I.1 shows the overlap that occurs during the simultaneous fielding of four panels in the fall round.

Table I.1 shows the number of Round 46 respondents by oldest age attained in 2006. Differential sampling rates were used to obtain the relatively large samples of the disabled and the oldest subgroups.

Complex Sample Design. In order to minimize survey costs while maximizing the precision of statistics, the sample was concentrated in over 1,200 ZIP code clusters within the 107 primary sampling units consisting of Metropolitan Statistical Areas and clusters of non-metropolitan counties. Although this clustering increases the sample size that can be afforded for a given budget, the precision is not as good as would be expected from a simple random sample of the same size (were such a sampling procedure feasible).

The use of differential sampling rates also has the effect of degrading precision compared to a simple random sample of the same size. Standard statistical packages, such as SAS, SPSS, S, and BMDP are not designed to control for clustering and differential sampling rates and can give seriously incorrect standard errors, confidence intervals and p-values. Two packages that will give correct estimates of precision and significance are WesVar and SUDAAN. See Section 5 for information on how to use these packages.

Table I.1 Number of completes at Round 46 by age category

Oldest age attained in 2006	Number of completes at Round 46
Total	15,770
Under 45	1,410
45 to 64	1,374
65 to 69	2,619
70 to 74	2,679
75 to 79	2,571
80 to 84	2,620
85 and older	2,497

Non-response Bias Knowing that cumulative attrition can become very serious in a panel survey, a concerted effort is constantly made to keep response rates high at each round. After conducting 10 rounds of interviewing, approximately 64 percent of the eligible sample beneficiaries in the 2003 Panel were still responding. (The other 36 percent dropped out of the survey after providing anywhere from 0 to 9 interviews.) See Table I.2 for a complete breakdown of cumulative response rates by panel for Rounds 37 through 46.

Although average cumulative non-response rates in this range are not trivial; steps have been and continue to be taken to reduce the risk of non-response bias. First, at each round, data from administrative records and from prior rounds are used to contrast new non-respondents with the continuing sample. Where systematic differences are noted, the survey sampling weights are adjusted to reduce potential biases. For this reason, it is critical that the survey sampling weights be used in all analyses. (This variable is called CS1YRWGT on the RIC X record format.)

Table I.2 Cumulative Response Rates for Each Panel Through Ten Rounds of Data Collection

Cumulative Response Rates for Medicare Current Beneficiary Survey by Round								
	2000 Panel Response Rate (n=6,037)	2001 Panel Response Rate (n=5,968)	2002 Panel Response Rate (n=5,967)	2003 Panel Response Rate (n=5,930)	2004 Panel Response Rate (n=5,978)	2005 Panel Response Rate (n=6,565)	2006 Panel Response Rate (n=6,675)	2006 NE Panel Response Rate (n=1,490)
Round 37	67.6%	68.0%	73.3%	83.2%				
Round 38	67.2%	66.7%	71.6%	77.3%				
Round 39	67.0%	65.8%	70.1%	73.9%				
Round 40		64.5%	68.0%	71.0%	82.2%			
Round 41		63.8%	66.9%	69.4%	76.3%			
Round 42		63.7%	65.9%	67.8%	73.2%			
Round 43			65.0%	66.6%	70.7%	82.0%		
Round 44			64.5%	65.4%	68.8%	75.1%		
Round 45			64.3%	64.6%	67.7%	72.1%		
Round 46				63.9%	66.3%	69.5%	82.8%	83.2%

Longitudinal Analyses. In addition to cross-sectional analyses of the Medicare population as of the fall of 2006 this PUF may be linked to PUFs from preceding years to enable longitudinal analyses. Special survey sampling weights are provided for the analysis of different time periods. Table I.3 shows the different possible periods that may be analyzed after linkage with the panels involved, sample sizes, and the appropriate weight. (It should be noted that the longitudinal files are not appropriate for doing certain types of analyses, such as mortality, disenrollment or characteristics of nonrespondents. If an analyst is interested specifically in analyzing such topics using the MCBS, they will need additional data.)

The Community Interview. Sample persons in the community (or appropriate proxy respondents) are interviewed using computer-assisted personal interviewing (CAPI) survey instruments installed on notebook-size portable computers. The CAPI program automatically guides the interviewer through the questions, records the answers, and compares them to edit specifications, thereby increasing the output of timely, clear, and high quality data. CAPI guides the interviewer through complex skip patterns and inserts follow-up questions where certain data were missing from the previous round's interview. When the interview is completed, CAPI allows the interviewer to transmit the data by telephone to the home office computer.

Table I.3 Possible longitudinal analyses using CY 2006 Access to Care

Period Covered	Last possible "accretion" (new entitlement) date	Earliest possible loss of entitlement	Applicable panels	Rounds Interviewed	Sample Size	Weight
Fall 2003 to Fall 2006	1/1/2003	Fall 2006	2003	R37, R40, R43 &, R46	3,131	L4YRSWGT
Fall 2004 to Fall 2006	1/1/2004	Fall 2006	2003, 2004	R40, R43 & R46	6,599	L3YRSWGT
Fall 2005 to Fall 2006	1/1/2005	Fall 2006	2003, 2004, 2005	R43 & R46	10,571	L2YRSWGT

These interviews yield a series of complementary data over time for each sample person on utilization of health services, medical care expenditures, health insurance coverage, sources of payment (public and private, including out-of-pocket payments), health status and functioning, and a variety of demographic and behavioral information (such as income, assets, living arrangements, family supports, and quality of life). Additionally, an access to care supplement is asked once a year in the September – December round.

An effort is made to interview the sampled person directly, but in case this person is unable to answer the questions, he or she is asked to designate a proxy respondent, usually a family member or close acquaintance. In Round 46 approximately 12 percent of the community interviews were done with proxies.

The Facility Interview. The MCBS conducts interviews for persons in long-term care facilities using a similar, but shortened instrument. A long-term care facility is defined as having three or more beds and providing long-term care services throughout the facility or in a separately identifiable unit. Types of facilities currently participating in the survey include nursing homes, retirement homes, domiciliary or personal care facilities, distinct long-term units in a hospital complex, mental health facilities and centers, assisted and foster care homes, and institutions for the mentally retarded and developmentally disabled.

If an institutionalized person returns to the community, a community interview is conducted. If he or she spent part of the reference period in the community and part in an institution, a separate interview is conducted for each period of time. Because of this, a beneficiary can be followed in and out of facilities, and a continuous record is maintained regardless of the location of the respondent.

The initial contact for the institutional interview is always with the facility administrator. Interviews are then conducted with the staff designated by the director as the most appropriate to answer each section of the questionnaire. It was decided early in the design of the study not to attempt interviews with the sample person or family members. The facility interview does not include attitudinal or other subjective items. The facility instruments include:

- (1) The Facility Screener - This instrument gathers information on the facility to determine the facility type. It is asked during the initial interview;
- (2) The Baseline Questionnaire - Gathers information on the health status, insurance coverage, residence history, and demographic items on supplemental sample beneficiaries in a facility setting and new admissions from the continuing sample. Selected information from this questionnaire is updated annually for continuing sample persons using an abbreviated version, The Facility Component Supplement to the Core Questionnaire; and
- (3) The Facility Core Questionnaire - Collects information on facility utilization, charge and payment information. This questionnaire is asked in every round but the initial one.

Data Linkage. MCBS interview data have been augmented with selected individual person-level administrative data (for example, buy-in status for Medicaid and Medicare capitation payments for managed care plan membership) and fee-for-service claims for Medicare-covered services. The addition of these data greatly enhances the analytic power of the survey-reported data alone. This results in a database which combines data that can only be obtained from personal interviews (survey data) with Medicare administrative data and Medicare claims data. All personal identifying information is removed to ensure confidentiality.

Design of the Access to Care Data File

The Access to Care data file is designed to provide early release of MCBS data related to Medicare beneficiaries' access to care. Rapid release of access data is achieved by omitting survey reported utilization and expenditure data. The claims information, while limited to program payments for covered services, third party payments for some Medicare secondary payer situations, and potential beneficiary liability, allows significant analysis of the impact of program changes on the beneficiary. This process eliminates the need for imputation of missing cost and payment variables and bypasses the reconciliation of the utilization and expenditure data collected in the survey with Medicare claims data.

The content of the Access to Care data file is governed by its central focus. In addition to questions from the access supplement concerning access to care, satisfaction with care and usual source of care, the file contains demographic and health insurance data and data on health status and

functioning. To facilitate analysis, the information collected in the survey is augmented with data on the use and program cost of Medicare services from Medicare claims data.

Contents of this Documentation

The rest of this manual contains detailed information about this public use file and specific background information intended to make the data more understandable. The sections are described below.

Section 1: Technical description of the public use file specifications and the structure of the public use file. It also provides a brief description and count of each of the record types in this file.

Section 2: Codebook of the file variables. This codebook is organized by record type and contains the question number (for data collected in the survey), and variable name, description and location in the record. Codes or possible values and value labels are also supplied. Frequencies for most variables (those with fewer than 120 distinct values) are also included in the codebook, as are notes concerning when variables are inapplicable (that is, questions were not asked due to skip patterns in the CAPI program). An index of variables is also included at the end of the codebook.

Variables in the CMS bill records are documented slightly differently. Record layouts are provided and are cross-walked to CMS data dictionary names. The data dictionary supplies a full explanation of all the variables and their various values.

Section 3: Notes on how individual variables were collected.

Section 4: Questionnaires used in Round 46. The questionnaires have been annotated with variable names to associate the questions with the codebook.

Section 5: A general description of the MCBS sample design, estimation procedures and projections. A brief discussion of response rates is also included. This section concludes with a comparison of the MCBS projections to CMS control figures.

References

Centers for Medicare and Medicaid Services, Office of the Actuary, published 2006 data from the National Health Accounts.

Medicare Current Beneficiary Survey

CY 2006 Access to Care

Section 1: File Structure

Data Specifications

The MCBS 2006 Access to Care file consists of a series of 41 separate files. Twenty of these files contain data on the MCBS sample persons; 20 files (READMEs) contain SAS code (SAS input statements, labels, and format statements) to facilitate the use of the data files by users who use SAS. The one remaining file contains SAS code used to produce a format library for all the data.

CMS releases the data for use with either mainframe or personal computer settings. Figure 1.1a and Figure 1.1b show file specifications such as file names, record counts and the associated README names.

Summary of the Data

The 20 data files represent completed Round 46 interviews with a sample of 15,770 Medicare beneficiaries, and supplemental information from the CMS Medicare files. Of these cases, 14,732 beneficiaries had community interviews and 1,038 beneficiaries had facility interviews.

Using the Data

All datasets are standard "flat" files to allow for processing with a wide variety of operating systems and programming languages. The datasets can be divided into two subject matter groups; files related to MCBS survey data with related Medicare administrative variables and files related to Medicare bill data.

There are 20 data files containing survey data and related summary administrative variables. For each of these files there is a "README" file which includes a SAS INPUT statement, a PROC FORMAT to interpret the coded fields, labels which provide more information about the variable than would be possible in an 8-character name, and a FORMAT statement which associates the code interpretations with the appropriate variables. CMS uses a single format library containing all formats used in all the READMEs. The program CMS uses to create a SAS format library contains all the formats and is included in this release.

Figure 1.1a File Specifications

File Name	Record Count	Associated Readme File Name
DATA\FLAT FILES\ RICK.dat	15,770	Data\Flat Files\Readme\rick.txt
DATA\FLAT FILES\ RICA.dat	15,770	Data\Flat Files\Readme\rica.txt
DATA\FLAT FILES\ RIC1.dat	15,770	Data\Flat Files\Readme\ric1.txt
DATA\FLAT FILES\ RIC2.dat	14,732	Data\Flat Files\Readme\ric2.txt
DATA\FLAT FILES\ RIC2F.dat	1,038	Data\Flat Files\Readme\ric2f.txt
DATA\FLAT FILES\ RIC2H.dat	10,091	Data\Flat Files\Readme\ric2h.txt
DATA\FLAT FILES\ RIC2P.dat	14,732	Data\Flat Files\Readme\ric2p.txt
DATA\FLAT FILES\ RIC3.dat	14,732	Data\Flat Files\Readme\ric3.txt
DATA\FLAT FILES\ RIC4.dat	15,770	Data\Flat Files\Readme\ric4.txt
DATA\FLAT FILES\ RIC5.dat	14,732	Data\Flat Files\Readme\ric5.txt
DATA\FLAT FILES\ RIC6.dat	1,038	Data\Flat Files\Readme\ric6.txt
DATA\FLAT FILES\ RIC7.dat	1,038	Data\Flat Files\Readme\ric7.txt
DATA\FLAT FILES\ RIC8.dat	15,770	Data\Flat Files\Readme\ric8.txt
DATA\FLAT FILES\ RICH.dat	2,286	Data\Flat Files\Readme\rich.txt
DATA\FLAT FILES\ RICKN.dat	14,947	Data\Flat Files\Readme\rickn.txt
DATA\FLAT FILES\ RICPD.dat	14,280	Data\Flat Files\Readme\ricpd.txt
DATA\FLAT FILES\ RICX.dat	15,770	Data\Flat Files\Readme\ricx.txt
DATA\FLAT FILES\ RICX2.dat	15,770	Data\Flat Files\Readme\ricx2.txt
DATA\FLAT FILES\ RICX3.dat	15,770	Data\Flat Files\Readme\ricx3.txt
DATA\FLAT FILES\ RICX4.dat	15,770	Data\Flat Files\Readme\ricx4.txt

Claims Files

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

Figure 1.1b Research Claims File Specifications

File Name	File Records	Associated Readme File Name
Data\Research Claims\Flat Files\HHA.dat	2,274	Data\Research Claims\Flat Files\Readme\readhha.txt
Data\Research Claims\Flat Files\HHArev.dat	52,463	Data\Research Claims\Flat Files\Readme\readhharev.txt
Data\Research Claims\Flat Files\HSP.dat	1,040	Data\Research Claims\Flat Files\Readme\readhsp.txt
Data\Research Claims\Flat Files\HSPrev.dat	4,865	Data\Research Claims\Flat Files\Readme\readhsprev.txt
Data\Research Claims\Flat Files\INP.dat	4,759	Data\Research Claims\Flat Files\Readme\readinp.txt
Data\Research Claims\Flat Files\INPrev.dat	74,097	Data\Research Claims\Flat Files\Readme\readinprev.txt
Data\Research Claims\Flat Files\OTP.dat	56,584	Data\Research Claims\Flat Files\Readme\readotp.txt
Data\Research Claims\Flat Files\OTPrev.dat	317,713	Data\Research Claims\Flat Files\Readme\readotprev.txt
Data\Research Claims\Flat Files\SNF.dat	1,720	Data\Research Claims\Flat Files\Readme\readsnf.txt
Data\Research Claims\Flat Files\SNFrev.dat	13,658	Data\Research Claims\Flat Files\Readme\readsnfrev.txt
Data\Research Claims\Flat Files\PHY.dat	587,684	Data\Research Claims\Flat Files\Readme\readphy.txt
Data\Research Claims\Flat Files\DME.dat	50,115	Data\Research Claims\Flat Files\Readme\readdme.txt

Note that the SAS input statements [README files] create one observation per data record for all of the MCBS files except the Physician/Supplier Claims and DME Claims. The SAS input statements for those claim types treat each line item as a separate observation with the claim-level detail repeating for each line item.

As an illustration of the structure of the README files, Figure 1.2 is a copy of the README file for the Survey Identification Record, RIC1.

Figure 1.2: Sample Text a README file (README.RIC1 illustrated)

```

INPUT
  @      RI C      $2.
  @      VERSI ON  $1.
  @      BASEI D   $8.
  @2     D_DCB     $6.
  @8     ROSTSEX    2.
  @20    D_AFEVER   2.
  @22    D_AFVI ET  2.
  @4     D_AFKORE   2.

```

Section 1: File Structure

@6	D_AFWWI	2.
@8	D_AFWW	2.
@0	D_AFGULF	2.
@2	D_AFI RAF	2.
@4	D_AFPEAC	2.
@6	D_NGEVER	2.
@8	D_NGALL	2.
@0	D_NGDSBL	2.
@2	D_VARATE	3.
@5	D_RACE2	2.
@7	HI SPORI G	2.
@9	SPCHNLNM	2.
@1	SPDEGRCV	2.
@3	SPMARSTA	2.
@5	I NCOME	3.
@8	I NCSRCE	\$2.
@0	D_DI VCUR	\$2.
@2	RACEAS	2.
@4	RACEAA	2.
@6	RACENH	2.
@8	RACEWH	2.
@0	RACEAI	2.
@2	RACEOTH	2.
@4	JOBSTAT	2. ;

LABEL	RI C	= "Record Identification Code"
	VERSI ON	= "Version Number"
	BASEI D	= "Unique SP Identification Number"
	D_AFEVER	= "SP ever served in armed forces (AF)?"
	D_AFKORE	= "SP served in AF during Korean conflict?"
	D_AFPEAC	= "SP served in AF during peace time?"
	D_AFVI ET	= "SP served in AF during Vietnam era?"
	D_AFWW	= "SP served in AF during World War I?"
	D_AFWWI	= "SP served in AF during World War II?"
	D_AFGULF	= "SP served in AF during Gulf War?"
	D_DI VCUR	= "Census division of SP's residence"
	D_AFI RAF	= "SP served in AF during Iraq/Af conflict?"
	D_DOB	= "Date of birth (YYYYMM)"
	D_NGEVER	= "SP ever active Nat'l Guard/ Reserve?"
	D_NGALL	= "All active duty spent in Nat'l Guard?"
	D_NGDSBL	= "Does SP have disability from service?"
	D_RACE2	= "Race of SP"
	D_VARATE	= "Current VA disability rating"
	I NCOME	= "Income range of SP"
	I NCSRCE	= "Source of SP income data"
	ROSTSEX	= "Sex of SP"
	SPCHNLNM	= "Number of children living"
	SPMARSTA	= "Marital status of SP"
	SPDEGRCV	= "Highest grade SP completed"
	RACEAA	= "Is SP Black or African-American?"
	RACEAI	= "Is SP American Indian or Alaskan Native?"
	RACEAS	= "Is SP Asian?"
	RACENH	= "Is SP Native Hawaiian/Pacific Islander?"
	RACEOTH	= "Is SP of another race?"
	RACEWH	= "Is SP Caucasian?"
	HI SPORI G	= "Is SP of Hispanic or Latino origin?"
	JOBSTAT	= "Is SP now working at job or business?"

FORMAT	
D_DOB	\$DTE6FMT.

Section 1: File Structure

D_VARATE	VADI SFMT.						
ROSTSEX	SEXFMT.						
D_DI VCUR	\$CENSUS.						
D_RACE2	RACE2FMT.						
HI SPORI G	YES2FMT.						
SPMARSTA	MARFMT.						
I NCOME	I NCFMT.						
I NCSRCE	\$FLGFMT.						
SPCHNLNM	CHI LDFMT.						
SPDEGRCV	DEGREFMT.						
RACEAS	RACEAA	RACENH	RACEWH	RACEAI	RACEOTH	I ND2FMT.	
D_AFEVER	D_NGEVER	D_NGALL	D_NGDSBL	D_AFVI ET	D_AFKORE	D_AFWW I	
D_AFWW	D_AFPEAC	JOBSTAT	D_AFGULF	D_AFI RAF	YES1FMT. ;		

Structure of the MCBS public use file(s)

As mentioned above, the data files can be divided into two subject matter groups: files containing survey data with related Medicare administrative variables and files containing Medicare bill data.

There are 20 data files in the survey and administrative summary data group:

- Key
- Administrative Identification
- Survey Identification
- Survey Health Status and Functioning (Community only)
- Survey Health Status and Functioning (Facility only)
- Survey Health Status and Functioning (Helper)
- Survey Health Status and Functioning (Prevention)
- Survey Access to Care
- Survey Health Insurance
- Survey Enumeration
- Survey Facility Residence History
- Survey Facility Identification
- Survey Interview
- Survey HMO Supplement
- Survey KN Supplement (Beneficiary Knowledge and Information Needs)
- Survey PD Supplement (Prescription Drug)
- Survey Cross-Sectional Weights
- Survey Longitudinal Weights (for individuals from the Round 37 Panel who completed Round 37, Round 40, Round 43 and Round 46 interviews)
- Survey Longitudinal Weights (for individuals from the Round 37 or Round 40 panel who completed Round 40 Round 43 and Round 46 interviews)
- Survey Longitudinal Weights (for individuals from the Round 37, Round 40 or Round 43 panel who completed Round 43 and Round 46 interviews)

There are seven types of Medicare bill records in the detailed utilization portion of the file:

- Inpatient hospital
- Skilled nursing facility
- Hospice
- Home health
- Outpatient
- Physician/supplier (Part B)
- Durable medical equipment

The bill records represent services provided during calendar year 2006 and processed by the CMS in conjunction with our administrative functions. To facilitate analysis, the Administrative Identification record contains a summary of the utilization that these bills present in detail.

All MCBS records begin with the same three variables: a record identification code (RIC), the version of the file (VERSION) and a unique number that identifies the person who was sampled (BASEID). These elements serve to identify the type of record and to provide a link to other types of records. To obtain complete survey information for an individual, an analyst must link together records for that individual from the various data files using the variable BASEID. In Round 46, none of the sample people has a record on every data file. Figure 1.3 provides an overview of the presence of data records on the various data files for community and facility respondents.

The tables that follow Figure 1.3 describe all of the types of records in this release – Table 1.A describes the survey and administrative records and Table 1.B describes the bill records.

Section 1: File Structure

Figure 1.3 The number of records present on each of the data files for community and facility sample respondents

Data files	Community respondents	Facility respondents
RIC K – Key record	1 per respondent	1 per respondent
RIC A – Administrative Identification	1 per respondent	1 per respondent
RIC 1 – Survey Identification	1 per respondent	1 per respondent
RIC 2 – Health Status and Functioning (community)	1 per respondent	none
RIC 2F – Health Status and Functioning (facility)	none	1 per respondent
RIC 2H – Health Status and Functioning (helper)	1, several, or none per respondent	none
RIC 2P – Health Status and Functioning (prevention)	1 per respondent	none
RIC 3 – Access to Care	1 per respondent	none
RIC 4 – Health Insurance	1 per respondent	1 per respondent
RIC 5 – Enumeration	1 per respondent	none
RIC 6 – Facility Residence History	none	1 per respondent
RIC 7 – Facility Identification	none	1 per respondent
RIC 8 – Interview	1 per respondent	1 per respondent
RIC H – HMO Supplement	1 per respondent	1 per respondent
RIC KN – Beneficiary Knowledge & Information Needs	1 per respondent	none
RIC PD – Prescription Drug	1 per respondent	none
RIC X – Cross-sectional Weights	1 per respondent	1 per respondent
RIC X2 – 2 years' Longitudinal Weights	1 per respondent	1 per respondent
RIC X3 – 3 years' Longitudinal Weights	1 per respondent	1 per respondent
RIC X4 – 4 years' Longitudinal Weights	1 per respondent	1 per respondent
Hospital bills *	1, several, or none per respondent	
Skilled nursing facility bills *	1, several, or none per respondent	
Hospice bills *	1, several, or none per respondent	
Home health bills *	1, several, or none per respondent	
Outpatient bills *	1, several, or none per respondent	
Physician/supplier bills *	1, several, or none per respondent	
Durable medical equipment bills *	1, several, or none per respondent	

* *These bills are summarized in the Administrative Identification record (RIC A), but are provided for more detailed analysis. If the sample person used Medicare benefits, there will be one or many bills, of one or many types, depending on what types of services were used. If the sample person used no Medicare benefits of a certain type, there will be no bills of that type. If the sample person used no Medicare benefits at all, there will be no bills. The RIC A summary provides information about how many services of each type will be found in the bill record files.*

Table 1.A – File Overviews

File: KEY

RIC: K

Number of Records: 15,770 - 1 for each person who completed an interview

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

In addition to the BASEID, the KEY file contains the type of interview conducted and other variables for classifying the beneficiary.

File: ADMINISTRATIVE IDENTIFICATION

RIC: A

Number of records: 15,770 - 1 for each person who completed an interview

Description: The Administrative Identification file contains information about the sample person from administrative records maintained by the Centers for Medicare and Medicaid Services. It contains basic demographic information (date of birth, sex), insurance information (Medicare entitlement, Medicaid eligibility, HMO enrollment), and summarizes the sample person's Medicare utilization for 2006.

File: SURVEY IDENTIFICATION

RIC: 1

Number of records: 15,770 - 1 for each person who completed an interview

Description: The Survey Identification file contains demographic information collected in the survey. To some extent, it parallels the demographic information provided in the Administrative Identification file (date of birth and sex, for example). Demographic information that is not available in the CMS records, such as education, income and military service, are also present.

Table 1.A – File Overviews

File: SURVEY HEALTH STATUS AND FUNCTIONING—COMMUNITY

RIC: 2

Number of Records: 14,732 - 1 for each person who completed an interview

Description: The Survey Health Status & Functioning file contains data about the sample person's health, including: self-reported height and weight, a self-assessment of vision and hearing, use of preventive measures such as immunizations and mammograms, avoidable risk factors (smoking, e.g.) and a history of medical conditions. Standard measures – activities of daily living (ADLs) and instrumental activities of daily living (IADLs) – also appear in this file.

File: SURVEY HEALTH STATUS AND FUNCTIONING—FACILITY

RIC: 2F

Number of Records: 1,038 - 1 for each beneficiary for whom an interview was completed

Description: The Survey Health Status & Functioning file contains data about the sample person's health, including: self-reported height and weight, a self-assessment of vision and hearing, use of preventive measures such as immunizations and mammograms, avoidable risk factors (smoking, e.g.) and a history of medical conditions. Standard measures – activities of daily living (ADLs) and instrumental activities of daily living (IADLs) – also appear in this file.

Table 1.A – File Overviews

File: SURVEY HEALTH STATUS AND FUNCTIONING—HELPER

RIC: 2H

Number of Records: 10,091

Description: The Survey Health Status And Functioning (Helper) file contains information about those persons responsible for assisting respondents in performing their activities of daily living (ADLs) and/or instrumental activities of daily living (IADLs). The number of helpers, the helper's relationship to the respondent, and the types of ADLs and IADLs the helper assists the respondent in performing are contained in this file.

NOTE: The number of records reflects the number of persons identified as assisting the survey respondent in performing various ADLs and IADLs. Therefore, it is possible to have one, several, or no records per respondent.

File: SURVEY HEALTH STATUS AND FUNCTIONING—PREVENTION

RIC: 2P

Number of Records: 14,732

Description: The Survey Health Status and Functioning (Prevention) file contains data about screening, diagnosis, treatment, and attitudes toward these common preventative health areas of interest: mammogram, papsmear, prostate, diabetes, colon cancer, Flu and Pneumonia shots, blood pressure, osteoporosis, smoking, alcohol consumption, and leading an active lifestyle.

Table 1.A – File Overviews

File: SURVEY ACCESS TO CARE

RIC: 3

Number of Records: 14,732 - 1 for each community person who completed an interview

Description: The Access to Care file contains information from the Access to Care and Satisfaction with Care sections of the questionnaire. Sample people were asked general questions about their use of all types of medical services in 2006 and about their usual source of medical care. This file also contains the sample people's assessment of the quality of the medical care that they are receiving.

File: SURVEY HEALTH INSURANCE

RIC: 4

Number of Records: 15,770 - 1 for each person who completed an interview

Description: The Survey Health Insurance file summarizes current health insurance information provided by the sample people.

NOTE: To limit the size of the RIC 4 record, only 5 private health insurance policies are detailed. For individuals in the sample that had more than 5 private health insurance policies, the total in the summary indicator is correct, but the number of plans detailed is less than the total. After a comparison of two MCBS files revealed a deviation in the number of health insurance policies held by the survey population, the editing procedures in the Access to Care RIC 4 were changed to include only those health insurance plans that the survey participant is currently enrolled with.

Table 1.A – File Overviews

File: SURVEY ENUMERATION

RIC: 5

Number of Records: 14,732 - 1 for each person who completed a community interview

Description: The Enumeration file contains information about the sample person's household. It reflects the size of the household, and the age and relationship of the people in it.

File: SURVEY FACILITY RESIDENCE HISTORY

RIC: 6

Number of Records: 1,038 - 1 for each person who completed a facility interview

Description: The Facility Residence History file summarizes the sample person's stay(s) in the facility, providing information about the admission and some limited information about the sample person's living arrangement prior to admission.

NOTE: In converting the MCBS Facility questionnaire to CAPI, full advantage was taken of the work done by the Agency for Health Care Policy and Research in its development of the CAPI protocol for the National Nursing Home Expenditure Survey (NNHES). The NNHES closely resembles the MCBS facility questionnaire in design and content. Adaptation was done carefully to insure both the continuity of the MCBS data and their comparability with NNHES. Consequently, both the RIC 6 and RIC 7 were modified. A more complete discussion of these modifications can be found in Section 3: Notes on Using the Data.

Table 1.A – File Overviews

File: SURVEY FACILITY IDENTIFICATION

RIC: 7

Number of Records: 1,038 - 1 for each sample person interviewed in a facility

Description: The Facility Identification file provides general characteristics of the institutions, most of the information from the facility screener. In several cases, more than one sample person resided in the same facility. In these cases the RIC 7 records are redundant (containing all of the same information), and differ only in the BASEID.

File: SURVEY INTERVIEW

RIC: 8

Number of Records: 15,770 - 1 for each person who completed an interview

Description: The Survey Interview file summarizes the characteristics of the interview, including type of questionnaire, duration, and whether or not the interview was conducted with a proxy respondent.

File: SURVEY HMO SUPPLEMENT

RIC: H

Number of Records: 2,286 - 1 for each sample person

Description: The HMO Supplement file augments information from the Access to Care and Satisfaction with Care sections of the questionnaire. Sample people who were currently enrolled in a State licensed HMO at the time of the interview were asked general questions about their health plans, to include access to and satisfaction with medical services in 2006. This file also contains the sample people's assessment of the quality of the medical care that they are receiving, types of additional coverage offered, and any out of pocket costs associated with the health plan.

Table 1.A – File Overviews

File: BENEFICIARY KNOWLEDGE AND INFORMATION NEEDS SUPPLEMENT

RIC: KN

Number of Records: 14,947 - 1 for each sample person

Description: This supplement combines new questions with some questions previously asked in the BK and BN supplements. Some of the original questions resulted from the need to measure beneficiaries' knowledge about new options to the Medicare program in response to the Balanced Budget Act of 1997. The data collected in this supplement will allow an evaluation of the impact of existing education initiatives by CMS. The KN section will also help to refine future CMS education initiatives by asking about information that beneficiaries may need, preferred sources for this information, and beneficiaries' access to insurance information. This data also presents the knowledge beneficiaries have gained from CMS publications.

NOTE: The KN Supplement was conducted in the January through April 2007 interview period. As a result, there is a discrepancy in the number of records. This record count variance is attributed to the retiring of a panel and natural attrition.

File: PRESCRIPTION DRUG SUPPLEMENT

RIC: PD

Number of Records: 14,280 - 1 for each sample person

Description: This supplement collects data on exposure to and knowledge of the new Medicare prescription drug benefit by the respondent. Questions probe for general knowledge and experiences with this program.

NOTE: The PD Supplement was conducted in the January through April 2007 interview period. As a result, there is a discrepancy in the number of records. This record count variance is attributed to the retiring of a panel and natural attrition.

Table 1.A – File Overviews

File: SURVEY CROSS-SECTIONAL WEIGHTS

RIC: X

Number of Records: 15,770 - 1 for each sample person

Description: The Cross-Sectional Weights file provides cross-sectional weights, including general-purpose weights and a series of replicate weights.

File: TWO-YEAR SURVEY LONGITUDINAL WEIGHTS

RIC: X2

Number of Records: 15,770 - 1 non-zero weight for each sample person who completed an interview in the current and first preceding years.

Description: The Longitudinal Weights file provides longitudinal weights, including general-purpose weights and a series of replicate weights.

File: THREE-YEAR SURVEY LONGITUDINAL WEIGHTS

RIC: X3

Number of Records: 15,770 - 1 non-zero weight for each sample person who completed an interview in the current and both of the two preceding years.

Description: The Longitudinal Weights file provides longitudinal weights, including general-purpose weights and a series of replicate weights.

File: FOUR-YEAR SURVEY LONGITUDINAL WEIGHTS

RIC: X4

Number of Records: 15,770 - 1 non-zero weight for each sample person who completed an interview in the current and all of the three preceding years.

Description: The Longitudinal Weights file provides longitudinal weights, including general-purpose weights and a series of replicate weights.

Table 1.A – File Overviews

File: HOSPITAL BILL

RIC: INP

Number of Records: 4,759

Description: Inpatient hospital bills for the MCBS population. These include bills from short stay general hospitals, and long-term hospitals such as psychiatric and TB hospitals. Different provider types are distinguishable. Generally, there is one bill for each stay. Some hospitals, particularly the long-term facilities, may bill on a cyclical basis and several bills may constitute a single hospitalization.

File: SKILLED NURSING FACILITY BILL

RIC: SNF

Number of Records: 1,720

Description: Skilled-nursing facility bills for the MCBS population. These include Christian Science facilities and other skilled nursing facilities. Different provider types are distinguishable. Generally, several bills constitute a period of institutionalization.

File: HOSPICE BILL

RIC: HSP

Number of Records: 1,040

Description: Hospice bills for the MCBS population. Billing practices vary by provider in that some hospices bill on a cycle (e.g. monthly) so that several bills constitute a period of hospice care; others submit a series of "final" bills.

Table 1.A – File Overviews

File: HOME HEALTH AGENCY BILL

RIC: HHA

Number of Records: 2,274

Description: Home health bills for the MCBS population. Home health agencies generally bill on a cycle, e.g., monthly.

File: OUTPATIENT BILL

RIC: OTP

Number of Records: 56,584

Description: Outpatient hospital bills for the MCBS population. These bills are generally for Part B services that are delivered through the outpatient department of a hospital (traditionally, a Part A provider).

File: PHYSICIAN/SUPPLIER BILL

RIC: PHY

Number of Records: 587,684

Description: Medicare Part B (physician, other practitioners, and suppliers other than DME--see RIC M below for DME) claims for the MCBS population. These records reflect services such as doctor visits, laboratory tests, X-rays and other types of radiological tests, surgeries, and inoculations.

File: DURABLE MEDICAL EQUIPMENT BILL

RIC: DME

Number of Records: 50,115

Description: Medicare DME Part B claims for the MCBS population. These records reflect claims for DME rentals and purchases.

Medicare Current Beneficiary Survey CY 2006 Access to Care

Section 2: Codebook

This release consists of two parts: 1) a summary segment, which contains all of the survey information and summary data from CMS's administrative and claims files; and 2) a bill segment, which contains itemized bill records from CMS' National Claims History (NCH) database.

The first part of this section, the Codebook, includes frequency counts for all of the variables in the summary segment. The second part of this section documents the variables (without frequencies) in the bill detail records.

SUMMARY SEGMENT

Using the Codebook

The Codebook lists of variables in each of the records, with their physical location in the record, their possible values and the question number from the questionnaire, if any. There are also some notes associated with most variables. Many questions are not always asked; the notes explain when questions are skipped. CMS also identifies when variables were first available if they were added after the initial MCBS year of 1991.

The first part of the Medicare Current Beneficiary Survey data file (that is, the survey and CMS summary data) is made up of 20 different types of records. The name of the record type being described is identified by name in the header. The RIC or record identification code with the record type being described is shown in the header on the first line on the right of the page above the page number. From time to time, CMS changes the files and as that happens, the version number changes. The version number appears in the header on the third line on the right of the page below the page number.

Variable - This column contains the variable names that we have associated with the SAS version of our data files. Since SAS limits variable names to 8 characters, these names are not always immediately meaningful. You can change them to more informative names, but the names in the tables were used to annotate the copies of the questionnaires.

Certain conventions apply to the variable names. All variables that are preceded by the characters "D_", such as D_SMPTYP are derived variables. The variables did not come directly from the survey data, but compiled from two or more survey variables. Variables preceded by the characters "H_", such as H_DOB, come for CMS source files.

Col (Column) - This column locates the variable physically in the record.

Len (Length) - This column describes the length of the field of the variable.

Fmt (Format) Name - This column identifies the format name associated with the variable in the SAS README file for this variable's RIC.

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

Ques # - The column headed "Ques #" contains a reference to the questionnaire for direct variables, or to the source of derived variables. For example, the "Ques #" entry that accompanies the variable ERVISIT in the Access to Care record is "AC1." The first question in the Access to Care portion of the community questionnaire is the one referenced. This column will be blank for variables that relate to neither the questionnaire or to CMS source files. These variables, such as the record identification code (variable name is RIC), are usually ones that we created to manage the data and the file.

Table 2.1 lists the abbreviations that may appear in this column when a section of the questionnaire is referenced.

Ty (Type) - This column identifies the type of variable, either numeric (N) or character (C).

Label (Variable label and codes) - In the first line under this column, you will find an explanation of the variable, which describes it more explicitly than would be possible in only 8 letters. These labels are available in README files, if you wish to use them in creating SAS data sets.

For coded variables all of the possible values of the variable appear in lines beneath that explanation. Associated with each possible value (in the column labeled "Frequency") is a count of the number of times that the variable had that value, and, under the column labeled "Label", a short format expanding on the coded value. Formats are also available in the README files.

Certain conventions were used in coding all variables to distinguish between questions that beneficiaries would not or could not answer and questions that were not asked. These conventional codes are as follows: "." or "-1" if the question was not applicable; "-7" if the respondent refused to answer; "-8" if the respondent didn't know the answer; and "-9" if the answer could not be ascertained from the response. With derived variables, a " " (blank) or "." mean that the variable could not be derived because one or more of the component parts was not available.

Many questions were posed to elicit simple "Yes" or "No" answers, or to limit responses to one choice from a list of categories. In these cases the responses are "Yes" or "No" or one of the codes from the list. In other questions the respondent was given a list of items to choose from and all of the responses were recorded. In these cases each of the responses is coded "Indicated" or "Not Indicated."

If a beneficiary responded with an answer that was not on the list of possible choices, it was recorded verbatim. All of the verbatim responses were reviewed and categorized. New codes were added to the original list of options to accommodate narratives that appeared frequently. For this reason the list of possible values for some variables may not exactly match the questionnaire.

Inapplicable - Each variable is followed by a statement that describes when a question was not asked, resulting in a missing variable. Questions were not asked when the response to a prior question or other information gathered earlier in the interview would make them inappropriate. For example, if the sample person said he has never smoked (community component, question HFG1), he would not be asked if he smokes now (question HFG2).

Table 2.1: Abbreviations Used to Identify Sections of the Questionnaires

Community Baseline Questionnaire

IN	Introduction	US	Usual Source of Care
EN	Enumeration	HF	Health Status and Functioning
HI	Health Insurance	DI	Demographics/Income
AC	Access to Care	HA	Housing Characteristics
SC	Satisfaction with Care		

Community Special Supplemental Questionnaire

NE	New Enrollee Supplement (Collected in Round 46)
KN	Knowledge and Needs Supplement (Collected in Round 47)
PD	Prescription Drug Supplement (Collected in Round 47)
RX	Drug Coverage Supplement (Collected in Round 48)
IA	Income and Assets Supplement (Collected in Round 48)

Facility Questionnaire (Screener)

FAVERIF, FA, FB

Facility Baseline Questionnaire

RH	Residence History
BQ	Background History
IN	Health Insurance
HA	Health Status and Functioning

BILL DETAIL SEGMENT

Using the tables

The tables in the bill detail section describe the Medicare utilization files included on the data file. There are two sets of tables; they must be considered together in order to interpret the data in this segment.

- **FILE DESCRIPTIONS FOR MEDICARE CLAIMS** - These record layouts correspond to the seven Medicare utilization files on the data file(s). The inpatient hospital and SNF bill files are described in the same record layout even though they are in separate datasets.

NCH No. - The number associated with each variable in the data file bill records and CMS' Data Dictionary (discussed below). The NCH No. can be used to crosswalk from the bill record to the more detailed description in the dictionary.

Variable - The name we have assigned to the data element (variable). Names may be up to eight characters long and are mnemonic. The variable name links the record layout to the remainder of the bill detail documentation. This name is also the name that we have supplied in the "README" SAS INPUT statement and labels.

Type - The format of the data element, or variable. Singly occurring data fields may be numeric, character or packed-decimal.

Group items may appear more than once, depending on the information that is present in the bill. For example, if several surgical procedures were reported on the bill, each of them would appear as a separate group item. One surgical procedure would translate to a single group item. A counter shows how many of each trailer type are present. For example, the number of ICD-9-CM procedure code groups present on the claim would be indicated by the counter PROCCNT.

Length - The number of bytes physically occupied by the variable in the record.

Format - How the data should be interpreted. For example, date fields may be read as eight characters, interpreted as CCYYMMDD (two-digit century, followed by two-digit year of the century, followed by two-digit month, followed by the two-digit day of the month).

Description - A more complete explanation of what the variable contains. These descriptions can be assigned to variables with the SAS LABEL code that is provided in the "README" file.

- **DATA DICTIONARY** - These tables are maintained by CMS to describe their internal records. They contain standard definitions of the variables in this file and values for all coded variables. Some of the variables referenced in this dictionary do not appear in this file. We have deleted some fields to protect the privacy of those who are participating in the survey.

Note: At the time this file was released, CMS was using version (I) of the Medicare claims. It is important to look closely at the readme files for each of the claims records before attempting to merge the Medicare claims data with MCBS data.

Medicare Current Beneficiary Survey

CY 2006 Access to Care

Section 3: Notes on Using the Data

In an undertaking of this nature and magnitude, there are bound to be questions about how terms are defined operationally, how field procedures affect the data collection process, complicated skip patterns, and how some variables were derived. We have included this section to address these questions.

This section is a collection of information about various data fields present in this release. We have not attempted to present information on every survey data field; rather, we concentrated our efforts on data fields where we have something useful to introduce. We start with information which is relevant across the board (global information). We follow that with specific information on individual data fields, presented in the same sequence as the data fields appear in the codebook.

Global Information

Missing Values

We use various negative values to indicate the reason why some data are missing. For instance, for survey collected data, a value of -1 indicates that the variable is inapplicable. A variable is generally inapplicable because the question did not need to be asked. For example, we do not ask male sample persons about hysterectomies. CMS uses other codes to denote standard reasons why data are not available, such as -7 for “refused,” -8 for “don’t know,” and -9 for “not ascertained.”

Dates

The CMS-derived dates of birth and death include century indicators and are in the format CCYYMMDD (2-digit century, 2-digit year, 2-digit month and 2-digit day). Due to the manner in which the responses were given, these dates must be evaluated in parts because one or more of the parts may be missing. For example, a vague response about a particular date (such as, “I know it was in June of last year, but I’m not sure of the exact day”) would be coded “200609-8” (“20” for the century, “06” for the year, “09” for September, and the code “-8” for “Don’t know” for the day).

Narratives

Respondents were asked a number of open-ended questions. The respondents answered these questions in their own words, and interviewers recorded the responses verbatim. The interviewer was prohibited from paraphrasing or summarizing the respondents’ answers. However, this release does not contain narratives. Instead, we have assigned codes to similar responses to make analysis easier. Often there will be more than one answer to a single question. In these cases, we supply several variables, all of which contain categorized data.

Specific variables - Key Record (RIC K)

There is one record for each individual who completed an interview, either community (**INTERVU** = “C”) or facility (**INTERVU** = “F”).

A facility interview was conducted whenever a sample person was residing in a facility: 1) that contains three or more beds; 2) that is classified by the administrator as providing long-term care, and 3) whose physical structure allows long-term care residents of the facility to be separately identified from those of the institution as a whole. This broad definition allows analysis beyond traditional views of long-term care, that is, nursing home and related care homes having three or more beds and providing either skilled nursing, or rehabilitative or personal care (other than supervision). Analysts can narrow or broaden the focus of their studies of facility care by using information from the Survey Facility Identification Record (RIC 7.) This file contains one record for each sample person for whom a facility questionnaire was administered.

This release is a mix of people who joined the survey in Round 37 (**D_SMPTYP** = “01”), Round 40 (**D_SMPTYP** = “02”), Round 43 (**D_SMPTYP** = “03”), or Round 46 (**D_SMPTYP** = “04”). CMS conducted interviews for the continuing sample (that is, the supplemental samples of Rounds 37, 40, and 43) in as many settings as necessary, to create a seamless view of the entire round. For the Round 46 supplemental sample, CMS conducted only one interview (facility or community) depending on the sample person’s living situation. A very small number of sample persons had more than one interview in a round as they moved to or from the community setting during a round. To avoid duplication of data, CMS only uses data obtained from the last interview in Round 46 for each sample person.

This file contains data from sample persons who were enrolled Medicare on January 1 and were alive for the fall round interview. This file is well suited for making estimates of the Medicare population who were enrolled for the entire year (the “always enrolled”). We estimate the “always enrolled” Medicare population in 2006 at 40.7 million beneficiaries. This group can be distinguished from all Round 46 interviews by selecting only beneficiaries who were enrolled before 2006 and survived until 2007 (**SURVIVE** = “Y”). It is inappropriate to use this file for estimates of either the “ever-enrolled” 2006 Medicare population or a “point-in-time” Medicare population because the sample excludes beneficiaries who either (1) were new to Medicare in 2006, or (2) died in 2006 prior to the fall round.

Administrative Identification Record (RIC A)

Except as noted otherwise, the variables in this record were derived from CMS’s Medicare enrollment database. History records were searched to establish the beneficiary's status (for example, age, residence, and type of beneficiary) as of December 31, 2006.

Section 3: Notes on Using the Data

The MCBS furnishes five variables relating to the sample person's age: four in the Administrative Identification record (RIC A), and one in the Survey Identification record (RIC 1). The “legal” dates of birth and death from Medicare and Social Security Administration records are recorded in RIC A (**H_DOB** and **H_DOD**, respectively.) [Note: the SSA legal date of birth is the first day of the birth month, unless the sample person was born on the first of the month, in which case the legal birth date is the first day of the previous month. Similarly the date of death (**H_DOD**) is the last day of the death month.] The variable **H_AGE** represents the sample person's “legal” age as of December 31, 2006. The variable **D_STRAT** groups the sample persons by various age categories using **H_AGE**. The sample person's date of birth, as reported by the respondent during the initial interview, is recorded in the RIC 1 (**D_DOB**).

During CY2006 an average of 7.1 million Medicare enrollees or 17.4 percent of the total population enrolled at that time had their Part B premium paid for by a State agency. (These data do not include those entitled to Medicaid but were not buy-ins; that is, the medically needy. This data also says nothing of how extensive the Medicaid benefits were.) This process, called State buy-in, is tracked by CMS and is used as a general proxy for Medicaid participation. CMS derived the variables that describe this participation (**H_MCSW**) and (**H_MCDE01 - H_MCDE12**) using its enrollment database. The variable (**H_MCSW**) can be used when only an indication that the enrollee was a “buy-in” at any time during 2005 is needed for analysis. The monthly variables (**H_MCDE01 - H_MCDE12**) can be used for analyzing Medicaid eligibility at specific points in time.

Some of the beneficiaries in the MCBS sample belong to Medicare managed care plans. CMS derived variables that describe this Medicare managed care membership (**H_GHPSW**) and (**H_PLTP01 - H_PLTP12**) were its enrollment database. The variable (**H_GHPSW**) can be used when only an indication that the enrollee was a member of a Medicare managed care plan at some time during 2006 is needed for analysis. The monthly variables (**H_PLTP01 - H_PLTP12**) can be used for analyzing membership at specific points in time.

Utilization Summary

For easier comparison of groups of people by the number and cost of medical services they have received, the Administrative Identification Record also includes a summary of all Medicare bills and claims for calendar year 2006, as received and processed by CMS through July 2007. (See the variables in the Administrative Identification Record from (**H_LATDCH**) to the end). In response to privacy regulations of the Health Insurance Portability and Accountability Act of 1996 (HIPAA), CMS has removed all “facial identifiers” (name, social security number, etc.) from the Medicare claims data. Researchers who need the “facial identifiers” to conduct their research must justify this need in their data request.

The utilization summary represents services rendered and reimbursed under fee-for-service in calendar year 2006. If a beneficiary used no Medicare services at all or was a member of a coordinated or managed care plan (such as a risk HMO) that does not submit claims to a fiscal intermediary or carrier, all program payment summary variables will be empty. If the beneficiary

Section 3: Notes on Using the Data

used no services of a particular type (for example, inpatient hospitalization), the variables relating to those benefits will be empty. Empty variables are zero-filled, except as noted in the next paragraphs.

Adjustment bills Initial claims submitted by fiscal intermediaries and carriers for services rendered and paid for by Medicare may be modified by later transactions that result in additional submittal of information relevant to payment or utilization for a given event. There are two types of Part A (institutional) adjustment transactions: credit-debit pairs, and cancel-only credit transactions. Both types of transactions cancel out a bill that was processed earlier (the credit bill exactly matches the earlier bill, which can be viewed as an initial debit). The difference between them lies in how (or if) a new debit transaction is applied to show the correct utilization. If the adjustment consists of a credit-debit pair, the new debit is applied immediately because it is submitted as the “debit” half of the pair. If the adjustment is a cancel-only transaction, the debit may be processed at a later date through a separate bill. In some cases, as when the original bill was completely in error, the cancel-only transaction simply serves to “erase” a mistake, and no new debit would be submitted. For this file, the adjustment processing removes the original debit and the credit that cancels it out, leaving only the final, corrected debit.

[NOTE: A few rare cases of credit bills with no prior debit may be in this file; these records can be dropped from analysis because they are, in effect, canceling out something for which CMS has no record.]

For Part B claims, we summarized only accepted claims (process code is “A”), or adjusted claims if the adjustment concerned money (process code either “R” or “S” and allowed charges greater than \$0). If the claim disposition code (DISPCD) was “03” or “63” (indicating a credit), both the credit and the matching debit were deleted.

Individual fields After adjustments were processed; the bills were summarized following the rules set forth below.

Inpatient hospital bills

Utilization is summarized by admissions, days, charges, covered charges, reimbursement amount, coinsurance days, and coinsurance amount. CMS calculated the number of admissions (**H_INPSTY**) by sorting the bills in chronological order, and counting the admission in each uninterrupted series of inpatient days. Total covered days (**H_INPDAY**) were summed from COVDAY in the bill. Total coinsurance days (**H_INPCDY**) were summed from COINDAY. Total bill charges and non-covered charges were selected from the revenue center trailer coded “0001”; total charges were summed as (**H_INPCHG**) and covered charges (total charges less non-covered charges) were summed as (**H_INPCCH**). Coinsurance amounts (**H_INPCAM**) were summed from COINAMTA in the bill. Reimbursement (**H_INPRMB**) is the sum of PROVPAY, organ acquisition costs (if any) and “pass through” amounts. Organ acquisition costs were accumulated from revenue center trailers when the second and third positions of the code were “81.” Pass through amounts

were calculated by multiplying covered days (COVDAY in the bill record) by the pass through per diem (PTDIEM in the bill record).

Skilled Nursing Facility

Utilization is summarized by admissions, days, charges, covered charges, reimbursement amount, coinsurance days and coinsurance amount. CMS calculated the number of admissions (**H_SNFSTY**) by sorting the bills in chronological order, and counting the admission in each uninterrupted series of inpatient days. Total covered days (**H_SNFDAY**) were summed from COVDAY in the bill. Total coinsurance days (**H_SNFCDY**) were summed from COINDAY. Total bill charges and non-covered charges were selected from the revenue center trailer coded "0001"; total charges were summed as (**H_SNFCHG**) and covered charges (total charges less non-covered charges) were summed as (**H_SNFCCH**). Total coinsurance amounts (**H_SNFCAM**) were summed from COINAMTA in the bill. Total reimbursement (**H_SNFRMB**) is the sum of PROVPAY.

Home Health

Utilization is summarized by the number of visits; visit charges; and other (that is, nonvisit) charges. If the second and third positions of the revenue center code were 42, 43, 44, 47, 55, 56, 57, or 58, then the units in the trailer (visits) were added to total visits (**H_HHAVST**) and the charges were accumulated as total covered visit charges (**H_HHACCH**). If the revenue center codes did not indicate visits, the charges were accumulated as other HHA charges (**H_HHACHO**). Total home health reimbursements were summed for Part A (**H_HHRMBA**) and Part B (**H_HHRMBB**).

Hospice

Utilization is summarized by days, covered charges and reimbursement amount. Covered hospice days (**H_HSDAYS**) were summed from the bill variable COVDAY. Covered charges were selected from the revenue center trailer coded "0001" and summed as (**H_HSTCHG**). Total hospice reimbursement (**H_HSREIM**) was summed from the variable PROVPAY.

Outpatient

CMS summarizes utilization by the number of bills, covered charges, and reimbursement amount. Total bills were counted as (**H_OUTBIL**). Total covered charges were selected from the revenue center trailer coded "0001" and summed as (**H_OUTCHG**). Total outpatient reimbursement (**H_OUTRMB**) was summed from the variable PROVPAY.

Part B (Carrier) claims

Utilization is summarized by: the number of claims; the number of line items; submitted and allowed charges; reimbursement; the number of office visits; and office visit charges. All claims and individual line items (there may be up to 13 line items per claim) were counted and summed as (**H_PHYCLM**) and (**H_PHYLIN**). Submitted charges and allowed charges (**H_PHYSCH**) and (**H_PHYACH**) were summed from SUBCRG and ALLOWCRG in the bill. Total reimbursement for Part B claims (**H_PHYRMB**) was summed from the variable PAYAMT in the bill. Durable Medical Equipment charges are broken out from the physician charges, as follows: number of claims (**H_DMECLM**), number of DME line items (**H_DMELIN**), DME submitted charges (**H_DMESCH**), allowed charges (**H_DMEACH**), and total DME reimbursement (**H_DMERMB**).

Office visits and their charges are summed with other services (described above) and as separate categories (**H_PMTVST** and **H_PMTCHO**). We summed office visits and office visit charges separately for two reasons. An office visit is a universally understood measure of service use and access to medical care. It also is an accurate measure of levels of service use across separate groups, unlike charge or payment figures that vary depending on the services that have been performed. Office visits are identified by HCPCS codes in the series 90000-90090 and 99201-99215 in the Part B line item trailer group(s).

Survey Identification Record (RIC 1)

"Initial interview" variables

Some questions are asked only in the initial interview of a sample person and are not asked again during subsequent sessions because the responses are not likely to change. Such questions include “Have you ever served in the armed forces?” and “What is the highest grade of school you ever completed?” Similarly, once a sample person has told us that he or she has a chronic condition (such as diabetes), the interviewer will not ask, “Have you ever been told you have diabetes?” in a subsequent interview. For this reason, the answers to these questions are missing from Round 46 for people from the 2003, 2004, and 2005 panels. To maximize the usefulness of this release as a cross-sectional file, we have these data forward from the initial interviews, for persons joining the survey in the 2003, 2004, and 2005 panels. Variables that have been reproduced this way are annotated “Initial interview” in this section.

When the complete date of birth was entered (**D_DOB**), the CAPI program automatically calculated the person's age, which was then verified with the respondent. In spite of this validation, the date of birth given by the respondent (**D_DOB**) does not always agree with the date of birth per SMS records (**H_DOB**). In these cases, the sample person was asked again, in the next interview, to provide a date of birth. Some recording errors have been identified this way, but in most cases beneficiaries provided the same date of birth both times they were asked. In some cases, proxies indicated that no one was exactly sure of the correct date of birth. In general, it is recommended that

Section 3: Notes on Using the Data

the variable (**H_DOB**) be used for analyses, since the CMS date of birth was used to select and stratify the sample. (Initial interview variable)

The VA disability rating (**D_VARATE**) is a percentage and is expressed in multiples of ten; it refers to disabilities that are officially recognized by the government as service-related. (Initial interview variable)

Race categories (**D_RACE2**) are recorded as interpreted by the respondent. Categories were not suggested by the interviewer, nor did the interviewer try to explain or define any of the groups. Ethnic groups such as Irish or Cuban were not recorded. (Initial interview variable)

Hispanic / Latino origin (**HISPORIG**) includes persons of Mexican, Puerto Rican, Cuban Central or South American or other Spanish culture or origin, regardless of race. Again, these answers are recorded as interpreted by the respondent. (Initial interview variable)

SPCHNLNM: Respondents were asked to report all living children, whether stepchildren, natural or adopted children. (Initial interview variable)

D_DIVCUR: The Census division is preformed through internal edits, by matching the survey participant's address to the appropriate Census region. The Census divisions are as follows:

New England – Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont;

Middle Atlantic – New Jersey, New York, Pennsylvania;

South Atlantic – Delaware, District of Columbia, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, West Virginia;

East North Central – Illinois, Indiana, Michigan, Ohio, Wisconsin;

West North Central – Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota;

East South Central – Alabama, Kentucky, Mississippi, Tennessee;

West South Central – Arkansas, Louisiana, Oklahoma, Texas;

Mountain – Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, Wyoming;

Pacific – Alaska, California, Hawaii, Oregon, Washington; and

Puerto Rico.

INCOME: Income represents the best source or estimate of income during 2006. Data gathered in Round 45 interviews represent the most detailed 2006 data and are used when available. For individuals not completing a Round 45 interview (that is, continuing panel people unavailable for that Round 45 and the Round 46 supplemental sample), the most recent information available was used. It should be noted that the variable **INCOME** includes income from all sources, such as pension, Social Security and retirement benefits, for the sample person and spouse. In some cases the respondent would not, or could not, provide specific information but did say the income was above or below \$25,000.

Section 3: Notes on Using the Data

Some “initial interview” questions (which were previously viewed as not likely to change) were carried over to subsequent sessions, because the responses were likely to change. The respondent was allowed to define marital status categories (**SPMARSTA**); there was no requirement for a legal arrangement (for example, separated).

Survey Health Status and Functioning Record – Community (RIC 2)

Part of the process of converting the facility instrument from a “paper and pencil” format to a Computer Assisted Personal Interviewing (CAPI) format was to adapt applicable questions from the facility instrument to the Resident Assessment Instrument (RAI) format. The RAI requires a Medicare and Medicaid certified long term care facility to conduct a comprehensive standardized assessment of the resident’s functional capacity and health status within 10 days of admission. In addition, a RAI must be completed once a year or whenever a resident’s health status changes. By adapting the applicable MCBS questions, interviewers can extract data regarding a resident’s health status and functioning directly from the RAI.

Note: Due to the number of variables that were altered in the facility instrument, resulting from the CAPI conversion, a separate RIC (2F) was created starting in 1997. As a result, RIC 2 now only includes data about the community population. The RIC 2F was created for the sample population responding to the health status and functioning section of the facility instrument.

The answers in the health status and functioning section of the questionnaire reflect the respondent's opinion, not a professional medical opinion.

Limitations on activities and social life (**HELMTACT**) reflect the sample person's experience over the preceding month, even if that experience was atypical.

For height and weight, the sample person was asked to recall or estimate, not to measure or weigh him or herself. In the height measurement (**HEIGHTFT** and **HEIGHTIN**), fractions of an inch have been rounded: those one half inch or more were rounded up to the next whole inch, those less than one half inch were rounded down. (Initial interview variable) In the weight measurement (**WEIGHT**), fractions of a pound have been rounded: those one half pound or more were rounded up to the next whole pound; those less than one half pound, were rounded down. (Initial interview variable)

Female SPs are asked if they have ever had a hysterectomy (**HYSTEREC**). "Hysterectomy" includes partial hysterectomies. (Initial interview variable).

Note: In subsequent rounds, if we learn that the SP has had a hysterectomy since the initial interview, the value for **HYSTEREC** is changed from 2 (“No”) to 1 (“Yes”)

Section 3: Notes on Using the Data

SPs are asked about whether they smoke. Use of other forms of tobacco, such as chewing tobacco, is not relevant to the “smoking” questions (**EVERSMOK** and **SMOKNOW**). Trying a cigarette once or twice was not considered “smoking”, but any period of regular smoking, no matter how brief or long ago, was considered smoking. We ask if the SP “Now” smokes. We define “Now” as being within the current month or so, and not just whether the sample person had a cigarette, cigar, or pipe tobacco on the day of the interview. Even the use of a very small amount at the present time qualified as a “yes.” Stopping temporarily (as for a cold) qualified as a “yes.” (**EVERSMOK** is an initial interview variable)

We ask SPs whether they have *ever* had any of a series of illnesses or conditions. Their responses were coded affirmatively if the sample person had at some time been diagnosed with the conditions, even if the condition had been corrected by time or treatment. The condition must have been diagnosed by a physician, and not by the sample person. Misdiagnosed conditions were not included. If the respondent was not sure about the definition of a condition, the interviewer offered no advice or information, but recorded the respondent's answer, verbatim. (Initial interview variables) We ask about: heart disease and high blood pressure (**OCARTERY**, **OCHBP**, **OCMYOCAR**, **OCCHD**, **OCCFAIL**, **OCCVALVE**, **OCRHYTHM**, and **OCOTHHRT**); disorders or diseases of the brain (**OCSTROKE**, **OCALZHR**, **OCPARKIN**); psychiatric disorders (**OCPSYCHO**); mental retardation (**OCMENTAL**); skin cancer (**OCCSKIN**); cancer, other than skin cancer (**OCCANCER**); diabetes (**OCBETES**); arthritis (**OCARTHRR** and **OCARTH**); osteoporosis (**OCOSTEOP**); a broken hip (**OCBRKHIP**); emphysema, asthma, or COPD (**OCEMPHY**); complete or partial paralysis (**OCPPARAL**); an amputation (**OCAMPUTE**); enlarged prostate or benign prostatic hypertrophy (**HAVEPROS**);

If the SP confirms having had cancer, *other than skin cancer* (**OCCANCER** = 1), we ask a series of follow-up questions to learn the affected body parts (**OCCLUNG**, **OCCCOLON**, **OCCBREAST**, **OCCUTER**, **OCCPROST**, **OCCBLAD**, **OCCOVARY**, **OCCSTOM**, **OCCCERVX**, **OCCKIDNY**, **OCCBRAIN**, **OCCTHROA**, **OCCBACK**, **OCCHHEAD**, **OCCFONEC**, **OCCOTHER**).

Has a doctor told you in the past year that you have a specific illness or condition?

All respondents are asked about various illnesses or conditions such as hypertension in the fall round. There are different versions of each question, depending on whether a respondent is in a supplemental sample (new panel) or continuing sample. New panel respondents are asked only if a doctor *ever* told them that they had a specific condition (hypertension, for example). Annually thereafter, the same respondents are asked if a doctor told them *in the past year* that they had a specific condition.

Since only those respondents in the continuing sample are asked the *in the past year* version of the question, CMS added a set of variables that are asked of new panel respondents. CMS combined responses to questions asked of the continuing sample with responses to questions asked of the

Section 3: Notes on Using the Data

supplemental sample. CMS did this so that there is one *in the past year* variable for all sample persons who completed a community interview in the fall round. Below is a list of these variables with the variables that CMS used to derived them:

Combined		Continuing Sample		Supplemental Sample
D_HBP	=	OCHBP	+	YRHBP
D_ARTHRD	=	OCARTHRD	+	YRARTHRD
D_BRKHIP	=	OCBRKHIP	+	YRBRKHIP
D_CANCER	=	OCCANCER	+	YRCANCER
D_CHD	=	OCCHD	+	YRCHD
D_CSKIN	=	OCCSKIN	+	YRSKIN
D_MYOCAR	=	OCMYOCAR	+	YRMYOCAR
D_PPARAL	=	OCPPARAL	+	YRPPARAL
D_PSYCHO	=	OCPSYCHO	+	YRPSYCHO
D_STROKE	=	OCSTROKE	+	YRSTROKE

D_PROST: Has a doctor told you in the past year that you had an enlarged prostate or benign prostatic hypertrophy (BPH)?

CMS combines the answers to two related questions on BPH in deriving the value for **D_PROST**. The first question (**HAVEPROS**) asks male SPs, who have not reported having had their prostate gland removed, if a doctor ever told them that they had BPH. This version of the BPH question is asked only once, in the very first MCBS interview. If an SP answers “Yes” to this question, a follow-up question is asked to determine if a doctor had told the SP that he had BPH in the past year (**YRPROST**). A different version of the same question (**HAVEPROS**) asks male SPs in the continuing sample, who have not reported having had their prostate gland removed, if a doctor told them in the past year that they had BPH. CMS combines the responses to **HAVEPROS** and **YRPROST** to yield a single variable, **D_PROST**.

Activities of Daily Living (ADL) and Instrumental Activities of Daily Living (IADL)

We ask SPs whether they have any difficulty performing 12 activities. Their answers about difficulty performing the IADLs (**PRBTELE**, **PRBLHWK**, **PRBHHWK**, **PRBMEAL**, **PRBSHOP**, and **PRBBILS**) ADLs (**HPPDBATH**, **HPPDDRES**, **HPPDEAT**, **HPPDCHAR**, **HPPDWALK**, and **HPPDTOIL**) and reflect whether or not the sample person usually had difficulty and anticipates continued trouble with these tasks, even if a short-term injury made them temporarily difficult.

Section 3: Notes on Using the Data

“Difficulty” in these questions has a qualified meaning. Only difficulties associated with a health or physical problem were considered. If a sample person only performed an activity with help from another person (including just needing to have the other person present while performing the activity), or did not perform the activity at all, then that person was deemed to have difficulty with the activity.

Help from another person includes a range of helping behaviors. The concept encompasses personal assistance in physically doing the activity, instruction, supervision, and “standby” help.

These questions were asked in the present tense; the difficulty may have been temporary or may be chronic. Vague or ambiguous answers, such as “Sometimes I have difficulty”, were coded “yes.”

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

The distinction between light housework (**PRBLHWK**) and heavy housework (**PRBHHWK**) was made clear by examples. Washing dishes, straightening up and light cleaning represent light housework; scrubbing floors and washing windows represent heavy housework. The interviewer was not permitted to interpret the answer in light of the degree of cleanliness of the dwelling.

PRBMEAL: “Preparing meals” includes the overall complex behavior of cutting up, mixing and cooking food. The amount of food prepared is not relevant, so long as it would be sufficient to sustain a person over time. Reheating food prepared by someone else does not qualify as “preparing meals.”

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

PRBBILS: Managing money refers to the overall complex process of paying bills, handling simple cash transactions, and generally keeping track of money coming in and money going out. It does not include managing investments, preparing tax forms, or handling other financial activities for which members of the general population often seek professional advice.

HPPDBATH: Those who have difficulty bathing or showering without help met at least one of the following criteria:

- someone else washes at least one part of the body;
- someone else helps the person get in or out of the tub or shower, or helps get water for a sponge bath;
- someone else gives verbal instruction, supervision, or stand-by help;
- the person uses special equipment such as hand rails or a seat in the shower stall;
- the person never bathes at all (a highly unlikely possibility); or,

- the person receives no help, uses no special equipment or aids, but acknowledges having difficulty.

HPPDDRES: Dressing is the overall complex behavior of getting clothes from closets and drawers and then putting the clothes on. Tying shoelaces is not considered part of dressing, but putting on socks or hose is. Special dressing equipment includes items such as button hooks, zipper pulls, long-handled shoe horns, tools for reaching, and any clothing made especially for accommodating a person's limitations in dressing, such as Velcro fasteners or snaps.

HPPDEAT: A person eats without help if he or she can get food from the plate into the mouth. A person who does not ingest food by mouth (that is, is fed by tube or intravenously) is not considered to eat at all. Special eating equipment includes such items as a special spoon that guides food into the mouth, a forked knife, a plate guard, or a hand splint.

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

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

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

D_ADLHNM: CMS derives the number of persons helping with ADLs and/or IADLs from HFK4a-f series and HFL9.

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

Survey Health Status and Functioning Record – Facility (RIC 2F)

CMS obtains much of the Health Status and Functioning data from the facility instrument to the Resident Assessment Instrument (RAI) format. The RAI requires a Medicare and Medicaid certified long term care facility to conduct a comprehensive standardized assessment of the resident's functional capacity and health status within 10 days of admission. In addition, a RAI must be completed once a year or whenever a resident's health status changes. By adapting the applicable MCBS questions, interviewers can extract data regarding a resident's health status and functioning directly from the RAI.

Note: Due to the number of variables that were altered in the facility instrument, resulting from the CAPI conversion, a separate RIC (2F) has been released since 1997. Previously, the data for SPs in facilities was included in RIC 2.

The answers in the health status and functioning section of the questionnaire reflect the respondent's opinion, not a professional medical opinion.

Limitations on activities and social life (**LIMACTIV**) reflect the sample person's experience over the preceding month, even if that experience was atypical.

For height and weight, the sample person was asked to recall or estimate, not to measure or weigh him or herself. In the height measurement (**HEIGHT**), fractions of an inch have been rounded: those one half inch or more were rounded up to the next whole inch, those less than one half inch were rounded down. (Initial interview variable) In the weight measurement (**WEIGHTF**), fractions of a pound have been rounded: those one half pound or more were rounded up to the next whole pound; those less than one half pound, were rounded down. (Initial interview variable)

Female SPs are asked if they have ever had a hysterectomy (**EVERHYST**). "Hysterectomy" includes partial hysterectomies. (Initial interview variable)

Note: In subsequent rounds, if we learn that the SP has had a hysterectomy since the initial interview, the value for **EVERHYST** is changed from 2 ("No") to 1 ("Yes")

HYSTLAST –This variable does not apply to and is not asked of:

1. male SPs;
2. female SPs in the new panel or supplemental sample other than those who reported that they have never had a hysterectomy (**EVERHYST** does not equal 2); or
3. female SPs in the continuing sample who previously reported having had a hysterectomy in an earlier round (**EVERHYST** = 1).

Survey participants are asked about whether they smoke. Use of other forms of tobacco, such as chewing tobacco, is not relevant to the "smoking" questions (**EVRSMOKE** and **NOWSMOKE**). Trying a cigarette once or twice was not considered "smoking", but any period of regular smoking,

Section 3: Notes on Using the Data

no matter how brief or long ago, was considered smoking. We ask if the SP “Now” smokes. We define “Now” as being within the current month or so, and not just whether the sample person had a cigarette, cigar, or pipe tobacco on the day of the interview. Even the use of a very small amount at the present time qualified as a “yes.” Stopping temporarily (as for a cold) qualified as a “yes.” (**EVRSMOKE** is an initial interview variable)

Activities of Daily Living (ADL) and Instrumental Activities of Daily Living (IADL)

We ask whether SPs have any difficulty performing 8 activities. Their answers about difficulty performing the ADLs (**PFBATHNG**, **PFDRSSNG**, **PFEATING**, **PFTRNSFR**, **PFLOCOMO**, and **PFTOILET**) and IADLs (**DIFUSEPH**, **DIFSHOP**, and **DIFMONEY**) reflect whether or not the sample person usually had difficulty and anticipates continued trouble with these tasks, even if a short-term injury made them temporarily difficult. Note that in addition to the three IADLs above that are common to both the community and facility interviews, the facility MDS evaluates five more IADLs (**IADSTOOP**, **IADLIFT**, **IADREACH**, **IADGRASP**, and **IADWALK**).

“Difficulty” in these questions has a qualified meaning. Only difficulties associated with a health or physical problem were considered. If a sample person only performed an activity with help from another person (including just needing to have the other person present while performing the activity), or did not perform the activity at all, then that person was deemed to have difficulty with the activity.

Help from another person includes a range of helping behaviors. The concept encompasses personal assistance in physically doing the activity, instruction, supervision, and “standby” help.

These questions were asked in the present tense; the difficulty may have been temporary or may be chronic. Vague or ambiguous answers, such as “Sometimes I have difficulty”, were coded “yes.”

PFBATHNG: Those who have difficulty bathing or showering without help met at least one of the following criteria:

- someone else washes at least one part of the body;
- someone else helps the person get in or out of the tub or shower, or helps get water for a sponge bath;
- someone else gives verbal instruction, supervision, or stand-by help;
- the person uses special equipment such as hand rails or a seat in the shower stall;
- the person never bathes at all (a highly unlikely possibility); or,
- the person receives no help, uses no special equipment or aids, but acknowledges having difficulty.

PFDRSSNG: Dressing is the overall complex behavior of getting clothes from closets and drawers and then putting the clothes on. Tying shoelaces is not considered part of dressing, but putting on socks or hose is. Special dressing equipment includes items such as button hooks, zipper pulls, long-

handled shoe horns, tools for reaching, and any clothing made especially for accommodating a person's limitations in dressing, such as Velcro fasteners or snaps.

PFEATING: A person eats without help if he or she can get food from the plate into the mouth. A person who does not ingest food by mouth (that is, is fed by tube or intravenously) is not considered to eat at all. Special eating equipment includes such items as a special spoon that guides food into the mouth, a forked knife, a plate guard, or a hand splint.

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

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

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

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

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

DIFMONEY: Managing money refers to the overall complex process of paying bills, handling simple cash transactions, and generally keeping track of money coming in and money going out. It does not include managing investments, preparing tax forms, or handling other financial activities for which members of the general population often seek professional advice.

Survey Health Status and Functioning Record – Helper (RIC 2H)

This file contains information about those persons responsible for assisting sample persons in performing their activities of daily living (ADLs) and/or instrumental activities of daily living (IADLs). The number of helpers, the helper’s relationship to the respondent, and the types of ADLs and IADLs the helper assists the respondent in performing are all contained in this file.

NOTE: The number of records reflects the number of persons identified as having assisted the sample person in performing one or more ADL or IADLs. Therefore, it is possible to have one, several, or no records per respondent.

HLPRREL – Relationship of Helper to SP: This is an automated function of the CAPI program. As people are identified as helpers with ADLs, IADLs, Medicare insurance decisions, or Medicare paperwork, their name is entered into the database. The respondent is then prompted for the helper(s) relationship to the SP. Both the House Hold Enumeration and Health Status sections provide names of individuals collected into this “name” and corresponding “relationship” database, found in one of three places:

1. proxy – **IN4a**;
2. anyone living with the SP – **EN2**; or
3. anyone who helped the SP with any personal care or daily needs – **HH21**.

HLPRNUM – Helper Identification Number: This variable is derived from the survey’s administrative files. The survey develops a person roster containing information about each person living with, treating or helping the SP. An internal variable, ROSTNUM, contains a sequential value assigned to each of these persons as they are identified by name during the interview. Individuals’ names may be gathered in four places:

1. proxy – **IN4a**;
2. anyone living with the SP – **EN1**;
3. any health or medical professional who treats the SP at home – **HH2**;
4. anyone who helped the SP at home with daily needs, but did not live with the SP, including home health aides, homemakers, friends, neighbors, or relatives – **HH19**.

The person roster number (contained in **ROSTNUM**) of a helper becomes the **HLPRNUM** for that helper.

Survey Health Status and Functioning Record – Prevention (RIC 2P)

Beginning in 2000, SPs were asked follow-up questions in regard to preventative services: mammogram, papsmear, prostate, diabetes, colon cancer, Flu and Pneumonia shots, blood pressure, osteoporosis, smoking, alcohol consumption, and leading an active lifestyle.

The following responses reflecting screening, diagnosis, treatment, and attitudes towards these common preventative health areas are rotated.

Preventative Service and the year the question was asked							
	YEAR						
	2000	2001	2002	2003	2004	2005	2006
Alcohol Consumption				✓		✓	
Blood Pressure				✓		✓	
Colon Cancer	✓			✓		✓	
Diabetes	✓		✓		✓		✓
Falls (Injuries/Concerns)							
Leading an active lifestyle		✓		✓		✓	
Osteoporosis	✓			✓			✓
Smoking		✓		✓		✓	

- **Smoking**—(DIDSMOKE, LASTSMOK, HAVSMOKE, QUITSMOK)
- **Alcohol consumption**—(DRINKDAY, DRINKSPD, FOURDRNK)
- **Leading an active lifestyle** – (D_VIGTIM, D_MODTIM, D_MUSTIM)
- **Colon cancer** – (COLHEAR, COLHTEST, COLHKIT, COLCARD, COLFDOC, COLRECNT, COLSCOPY, WHENSCOP, HEARSCOP, COLDREC, COLSCRNS)
- **Blood pressure**—(HYPEAGE, HYPEHOME, HYPEDRNK, HYPEMEDS, HYPELONG, HYPEMANY, HYPEPAY, HYPESKIP, HYPECOND, HYPECTRL)
- **Falls** – (FALLHELP, FALOINJ, FALLFRACT, FALSPRAN, FALBRUIS, FALCUT, FALCONC, FALDISLO, FALOTHER, FALLIMIT, FALLBACK, FALLFEAR)
- **Diabetes**—(DIAAGE, DIAPRGNT, DIAINSUL, DIAMEDS, DIATEST, DIASORES, DIAPRESS, DIAASPRN, INSUTAKE, INSUDAY, INSUWEEK, MEDSTAKE, MEDDAY, MEDWEEK, MEDMONTH, TESTTAKE, TESTDAY, TESTWEEK, TESTMNTH, TESTYEAR, SORECHEK, SOREDAY, SOREWEEK, SOREMNTN, SOREYEAR, DIATENYR, DIADRS AW, DIAHEMOC, DIACTRLD, DIAHYPO, DIAHYPTR, DIAFTEVR, DIAFEET, DIANEURO, DIACIRCF, DIAULCER, DIASKINC, DIAEYPRB, DIAKDPEV, DIAKDPRB, DIAKIDNY, DIAMNGE, DIATRAN, DIAKNOW, DIASUPPS, DIAEVERT, DIARECNT, DIAAWARE, DIARISK, DIASIGNS)
- **Osteoporosis**—(OSTEVERT, OSTHRISK, OSTFRACT, OSTTEST, OSTHEAR, OSTRECNT, OSTMASS)

Survey Access to Care Record (RIC 3)

Definitions applied to medical providers

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

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

Doctor's clinic practice – This refers to a group of doctors who have organized their practice in a clinic setting and work cooperatively; generally, patients either come in without an appointment or make an appointment and see whatever doctor is available.

HMO – This is an organization that provides a full range of health care coverage at a fixed periodic fixed fee, usually with low or no coinsurance requirement.

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

Free-standing surgical center - A facility performing minor surgical procedures on an outpatient basis, and not physically connected to a hospital.

Rural health clinic - provides outpatient services, routine diagnostic services for individuals residing in an area that is not urbanized and is designated as a health staff shortage area or an area with a shortage of personal health services. These services are provided for a nominal copayment and deductible.

Company clinic - A company doctor’s office or clinic which is operated principally for the employees (and sometimes their dependents).

Other clinic - a non-hospital facility such as a drug abuse clinic, a “free” clinic, a family planning clinic or military base clinic.

Walk-in urgent center - a facility not affiliated with a nearby hospital, offering services for acute conditions. Typically, people are seen without appointments.

Section 3: Notes on Using the Data

Home (doctor comes to sample person's home) - home is anywhere the sample person is staying; it may be his or her home, the home of a friend, a hotel room, etc.

Hospital emergency room - means the emergency room of a hospital. “Urgent care” centers are not included. (NOTE: All hospital emergency room visits were included, even if the sample person went there for a “non-emergency” condition such as a cold, flu or intestinal disorder.)

Hospital outpatient department - unit of a hospital, or a facility connected with a hospital, providing health and medical services to individuals who receive services from the hospital but do not require hospitalization.

Differences in the questionnaire sequence for the continuing and supplemental panel

It should be noted in using data in this section that the questionnaire sequence on access to care for supplemental panel persons differs from that for continuing panel persons and may lead to apparent differences in expected number of responses to questions in the access to care codebook section. For example, continuing panel persons indicating use of emergency room (and later, outpatient hospital) care in the utilization section of the core questionnaire are asked, after the conclusion of questions on utilization in that section, appropriate access to care questions about the visit (AC3-AC6). The CAPI program then reverts back to the next utilization section in the core questionnaire. Questions AC1 and AC2 are not later asked of these people.

The supplemental panel people, on the other hand, are not asked the core questions during their initial interview and go through the entire sequence of access to care questions. Thus, the number of persons responding to AC3 on whether or not they had an appointment (ERAPPT) is greater than those who responded to question AC1 on whether they had gone to a hospital emergency room for medical care during the reference period (ERVISIT).

Open-ended questions

Respondents were asked a number of open-ended questions (reasons for dissatisfaction with care, kinds of problems experienced in getting health care, etc.). The respondents answered these questions in their own words, and interviewers recorded the responses verbatim. The interviewer was prohibited from paraphrasing or summarizing the respondents' answer.

This file contains no verbatim responses. We have supplied, instead, codes that summarize the answer. Often there will be more than one code because the answer included several specific topics.

Other variables

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

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

USHOWLNG: If the sample person had an actual visit with the doctor listed in **USUALDOC** by the time of the interview, “less than one year” was coded.

OPDSCOND – (AC10): Was visit to OPD for a specific condition?

This question applies only to new panel SPs (RIC K, **D_SMPTYP** = 00) where either the reason for an OPD visit was not a medical condition (**OPDMCOND** is not equal to 1) OR the reason for an OPD visit was for a medical condition (**OPDMCOND** equal 1) but it was not for surgery (**OPDSURG** is not equal to 1).

MDSPCLTY – (AC20): What is the specialty of the physician who saw the SP most recently in a setting other than at home or in a hospital?

This question applies to:

1. all SPs in new panels; or
2. continuing SPs with no emergency room, outpatient department, or medical provider visits in either of the previous two rounds AND AC20, AC21, AC24-AC36 not already asked in this Round in the medical provider utilization section of the questionnaire.

MDSCOND – (AC22): Was visit to doctor’s office for a specific condition?

This question applies only to new panel SPs (RIC K, **D_SMPTYP** = 00) where either the reason for an OPD visit was not for a medical condition (**OPDMCOND** is not equal to 1) OR the reason for an OPD visit was for a medical condition (**OPDMCOND** equals 1) but it was not for surgery (**OPDSURG** is not equal to 1).

HEARMHMO – (HIMC1aa): Has the SP heard of an HMO that Medicare beneficiaries are able to join?

This question applies to initial interviews of all SPs in new panels and SPs in the continuing sample who have never been enrolled in a Medicare HMO.

Survey Health Insurance Record (RIC 4)

To help the respondent answer the questions about Medicaid, the interviewers used the name of the Medicaid program in the state where the sample person was living.

A health insurance plan is one that covers any part of hospital bills, doctor bills, or surgeon bills. It does not include any of the following:

- Public plans, including Medicare and Medicaid, mentioned elsewhere in the questionnaire.

Section 3: Notes on Using the Data

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

D_MCRHMO: What data source was used to determine the Medicare HMO enrollment stats? This variable was derived from CMS administrative records (RIC A: **H_ENT08 – H_ENT12**) as of the date of the interview in the fall round (RIC 8: **INT_DATE**.)

D_PRIVAT: What private health insurance coverage does the SP have, if any? This variable is derived from responses to **D_OBTNP1 – D_OBTNP5**.

D_MCARE: What Medicare coverage does the SP have? This variable is derived from CMS administrative records (RIC A: **H_ENT08 – H_ENT12**) data as of the date of the interview in the fall round (RIC 8: **INT_DATE**.)

D_MCAID: Is the SP eligible for Medicaid? This variable is derived from either CMS administrative records (RICA: **H_MCDE08 – H_MCDE12**) or survey data (**HIS7** or **HI7**) as of the date of the interview in the fall round (RIC 8: **INT_DATE**.)

MCAIDHMO: Was the SP enrolled in a Medicaid managed care plan? This variable was obtained from either (**HI10a**) or (**HIS10a**).

CHOICHMO: Was the SP given the opportunity to enroll in a Medicaid HMO? This variable was obtained from either (**HI10b**) or (**HIS10b**).

PUBRXCov: Does the SP's public plan cover prescription drugs? This variable was obtained from either (**HI16a**) or (**HIS16a**).

MCDRXCOV: Does the SP's Medicaid plan cover prescription drugs?

This variable was obtained from either **(HI10d)** or **(HIS10c)**.

D_HMOTYP: This variable is derived from CMS administrative records (RICA: **H_PLTP08 – H_PLTP12**) as of the date of the interview in the fall round (RIC 8: **INT_DATE.**)

D_HMOCOV: Was the SP enrolled in an HMO at any time in the last year?

This variable was derived from **(HIMC1a)**, **(HIMC1c)**, and **(HIMC1)**.

D_HMOCUR: Is the SP now enrolled in a Medicare HMO?

This variable was derived from either **(MC1)** or **(HIMC3)**.

D_DMCCOST: What is the annual cost of your drug discount membership cost?

This annualized cost was derived using the amount paid (**DMFEEAMT**) and the associated payment interval (either **DMFEEPAY** or **DMFEEOS**).

D_DMCMEM: How many active discount drug memberships does the SP have?

This variable is derived by adding the number of memberships from the last round (**DMCMHAVE**) to the number of new memberships entered into since the last round (**DMCMNEW**).

Detailed information is given for up to five health insurance plans in the plan trailer portion of the RIC 4. **D_TYPP1 – 5** is the plan type (private employer-sponsored insurance, private self-purchased insurance, unknown private insurance, private HMO or Medicare HMO). Note that private insurance plan information collected in the facility is categorized as 'unknown' because we do not ask the facility representative the source of the respondent's private health insurance.

D_PHREL1 – 5: What is the relationship of the policyholder to the SP?

The "Policy Holder" or "Main insured person" is the member of the group/union or the employee of the company that provides the insurance plans. It would also be the name on the policy, if the respondent had it available. Responses from **(HIS26)**, **(HI22a)**, or **(HI26)** are coupled with roster information which contains these relationships to determine the policyholder's relationship to the SP.

D_HMOPL1 – 5 indicates whether the plan is an HMO.

D_COVNM1 – 5: Is the number of people covered by each private plan.

This information is obtained from either: **(HI22d)**, **(HI29)** or **(HIS29)**.

D_COVRX1 – 5: Indicates whether the private plan covers prescription drugs.

This information is obtained from either: **(HI22e)**, **(HI30)**, or **(HIS30)**.

D_COVNH1 – 5: Indicates whether the private plan has long-term care coverage.

This information is obtained from either: **(HI22f)**, **(HI31)**, or **(HIS31)**.

D_PAYSP1 – 5: Does the main insured person (MIP) pay any part of the insurance premium? Obtained from either (HI22g), (HI32), or (HIS32).

D_ANAMT1 – 5: The annual cost of private health insurance plan premiums.

A premium amount was recorded even if the sample person did not directly pay the premium (if, for example, a son or daughter paid the premium). Premium amounts have been annualized based on the assumption that the SP held the policy for the entire 12-month period. This variable was derived using responses from (HI22h), (HI33) and (HIS33).

D_OBTNP1 – 5: How did the main insured person get the policy (e.g. self-purchased, employer, etc.)?

Obtained from either (HI22b), (HI27), or (HIS27).

D_INDUS1 – 5: The industry of the SP's current or former employer through which the SP obtained health insurance coverage.

Obtained from either (HI22c), (HI28), or (HIS28).

D_PLLTR1 – 5: What is the plan letter of the Medicare supplemental or Medigap policy? (Consists of a letter code of the ten standardized Medigap policies as reported by the respondent.)

Obtained from either (HI22b2), (HI27b) or (HIS27b).

D_INS1 – 5 Specifies whether the private health insurance plan has limited service coverage such as dental-only, prescription drug-only, etc. This information was developed through an editing process in which the plan names were researched and categorized into comprehensive insurance or single-service insurance plans. Furthermore, **D_RX1 – 5** was developed in conjunction with that editing process. **D_RX1 – 5** indicates if the private health insurance plan covers prescription drugs or is a prescription drug discount card. For example, a respondent may indicate that a plan covers drugs. If further analysis reveals that the plan is a single-service type, **D_RX1 – 5** would indicate no drug coverage in order to prevent drug imputation for all services. These flags were developed specifically to aid in accurately setting prescription drug imputation flags.

The following nine variables are found in both the RIC 4 and RIC H:

MHMORX: Does the SP have prescription drug coverage through the Medicare HMO? Obtained from either (HIMC6) or (HISMC4);

MHMODENT: Does the SP have dental coverage through the Medicare HMO? Obtained from either (HIMC7), (HISMC5), (HI22e1), or (HI30a);

MHMOEYE: Does the SP have optical coverage through the Medicare HMO?

Obtained from either (HISMC6), (HIMC8), (HI22e2), or (HI30b);

MHMOPCAR: Does the SP have preventive care coverage through the Medicare HMO?
Obtained from either (HISMC7), (HIMC9), (HI22e3), or (HI30c);

MHMONH: Does the SP have nursing home coverage through the Medicare HMO?
Obtained from either (HISMC8) or (HIMC10);

MHMOPAY: Besides the cost of the Medicare Part B premium, is there any additional cost for coverage, excluding co-payment amounts?
Obtained from either (HISMC9) or (HIMC11);

MHMOCOST: Does anyone else contribute to the cost of Medicare HMO coverage?
Obtained from either (HISMC11), (HIMC12a), (HI22h1), or (HI33a);

MHMOWHO: If anyone else did contribute to the cost of Medicare HMO coverage, who was it?
Obtained from either (HISMC12), (HIMC12b), (HI22h2), or (HI33b); and

D_ANHMO: What is the annual additional cost of Medicare HMO premiums?
The premiums have been annualized regardless of the length of time the sample person actively held the policy (HISMC10).

Survey Enumeration Record (RIC 5)

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

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

Because the date of birth or exact relationship of a household member was sometimes unknown (perhaps because a proxy provided the information), the sum of the variables "number related" / "number not related" (**D_HHREL/D_HHUNRL**) or "number under 50" / "number 50 or older"

(**D_HHLT50/D_HHGE50**) may not equal the total number of people in the household (**D_HHTOT**).

Survey Facility Residence History (RIC 6)

This file provides information on the sample person's admission to a facility. (ORIGADMN) gives the first admission date while (ADMIN) gives the most recent admission date. (LIVWRELA) tells with whom the SP was living prior to being admitted. (ADMTFROM) tells the place from where the SP was coming.

Survey Facility Identification Record (RIC 7)

This file provides general characteristics of the institutions, most of the information coming from the facility screener. Sometimes, more than one sample person resided in the same facility. In these cases the RIC 7 records are redundant (containing all of the same information), and differ only in the BASEID. There is one record for each sample person interviewed in a facility.

The value of variables representing “number of beds” (**FACLTBED** and **FACTOBED**) will be missing when either there were no beds of that type in the facility, or the question was skipped.

FACLTBED: The number of long-term beds is calculated from (**FA43**, **FA44**, **FA45**, **FA45a**, **FA45b**, **FA45c**, and **FA46**).

SPIDCNT: Backend edit count of survey participants that reside in a particular facility.

Survey Interview Description Record (RIC 8)

This record was added in the 1992 MCBS Access to Care public use release. Most of the material in it was included in the Survey Identification record in the 1991 MCBS Access to Care public use release.

Multiple Interviews

Some sample people had more than one interview in this round. To avoid duplication of data, the information in this file represents the last interview conducted with the sample person in Round 43. The variable **INTERVU** indicates which type of interview was conducted. Please see the description of the KEY Record (RIC K) earlier in this section for a more detailed description of multiple interviews and of this variable.

Proxy rules

Wherever possible, the community interviews were conducted directly with the sample person. In most cases, the sample person was able to respond to the interview unassisted. In a few cases, the sample person was assisted with the interview by a friend or relative, and in some cases the sample person was too ill or otherwise incapacitated to be interviewed. The variables **PROXY**, **D_PROXR**, **RRECHLP** and **D_IHLPR** provide information about who was interviewed, and how those respondents are related to the sample person.

People who were too ill, or who could not complete the community interview for other reasons were asked to designate a proxy, someone very knowledgeable about the sample person's health and living habits. In many cases, the proxy was a close relative such as the spouse, a son or daughter. In other cases, the proxy was a non-relative like a close friend or caregiver. The variable **PROXY** indicates whether or not a community interview was conducted with a proxy respondent, and the variable **D_PROXR** indicates the relationship of the proxy to the sample person. (Since all facility interviews are conducted with proxy respondents, this variable is "missing" for facility cases.)

If the sample person appeared confused or disoriented at the time of the interview, and no proxy could be identified, the interviewer was instructed to complete the questionnaire as well as possible. If the interviewer felt that the respondent was not able to supply reasonably accurate data, this perception was recorded in the interviewer remarks questionnaire and appears in this record as the variable **RINFOSAT**.

"Proxy needed – language problem" was given as a reason for the use of a proxy in less than 1 percent of the cases. More often, language problems were addressed without the use of a proxy. Interpreters were used in some cases, and bilingual interviewers used Spanish-language versions of the questionnaires when the respondent preferred to be interviewed in Spanish. There are both English and Spanish versions of the CAPI survey instrument; the variable **LANG** indicates which version was used.

Proxy respondents were always used in nursing homes, homes for the mentally retarded, and psychiatric hospitals. Sample persons were interviewed directly in prisons when that was permitted. The need for a proxy when interviewing respondents in other institutions was evaluated on a case-by-case basis.

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

Other variables

Several questionnaires are administered in the facility interview: a personal baseline for individuals
Medicare Current Beneficiary Survey

CY 2006 Access to Care

Section 3: Notes on Using the Data

in the supplemental sample found to reside in a nursing facility and for new admissions to a facility from the continuing sample; the core and supplement questionnaires for the continuing sample. The facility screener was administered in every case. Please see Section 4 for copies of all of the instruments and for a more detailed description of when each is administered.

Two variables are supplied to further characterize the interview: **LENGTH** contains the length of the interview, in minutes, and **RESTART** indicates whether or not the interview was interrupted. Community interviews are sometimes interrupted to accommodate the respondent's schedule or for other reasons. We did not calculate the duration of the community interview if the interview was interrupted. Facility interviews are conducted with several instruments and often involve a number of respondents. Since nearly all of the facility interviews are interrupted and total duration is difficult to capture (and interpret), **LENGTH** and **RESTART** are always missing for facility interviews.

INT_DATE: Date in which the interview was conducted.

Survey HMO Supplement Record (RIC H)

CMS added this record to the 1996 MCBS Access to Care public use release as a result of growing interest on the coverage and service provided by Medicare HMOs to Medicare beneficiaries. The questions in this record were asked only if the sample person was currently enrolled in a Medicare HMO.

The following variables are found in both the RIC 4 and RIC H codebooks:

MHMORX: Does the SP have prescription drug coverage through the Medicare HMO?
Obtained from either (**HIMC6**) or (**HISMC4**);

MHMODENT: Does the SP have dental coverage through the Medicare HMO?
Obtained from either (**HIMC7**), (**HISMC5**), (**HI22e1**), or (**HI30a**);

MHMOEYE: Does the SP have optical coverage through the Medicare HMO?
Obtained from either (**HISMC6**), (**HIMC8**), (**HI22e2**), or (**HI30b**);

MHMOPCAR: Does the SP have preventive care coverage through the Medicare HMO?
Obtained from either (**HISMC7**), (**HIMC9**), (**HI22e3**), or (**HI30c**);

MHMONH: Does the SP have nursing home coverage through the Medicare HMO?
Obtained from either (**HISMC8**) or (**HIMC10**);

MHMOPAY: Besides the cost of the Medicare Part B premium, is there any additional cost for coverage, excluding co-payment amounts?

Obtained from either (HISMC9) or (HIMC11);

MHMOCOST: Does anyone else contribute to the cost of Medicare HMO coverage?
Obtained from either (HISMC11), (HIMC12a), (HI22h1), or (HI33a);

MHMOWHO: If anyone else did contribute to the cost of Medicare HMO coverage, who was it?
Obtained from either (HISMC12), (HIMC12b), (HI22h2), or (HI33b); and

D_ANHMO: What is the annual additional cost of Medicare HMO premiums?
The premiums have been annualized regardless of the length of time the sample person actively held the policy (HISMC10).

Survey Beneficiary Knowledge and Information Needs Supplement (RIC KN)

This supplement was conducted during the January through April 2007 interview period, as opposed to the September through December (2006) interview period (with the rest of the Access to Care file). As a result of collecting this information during the following interview period, a reduction in sample size occurs. This reduction is due to normal attrition caused either by sample persons' death or refusal to continue their participation in the MCBS.

Survey Prescription Drug Supplement (RIC PD)

This supplement was conducted during the January through April 2007 interview period, as opposed to the September through December (2006) interview period (with the rest of the Access to Care file). As a result of collecting this information during the following interview period, a reduction in sample size occurs. This reduction is due to normal attrition caused either by sample persons' death or refusal to continue their participation in the MCBS.

Survey Cross-sectional Weights Record (RIC X)

Cross-sectional and three sets of longitudinal weights are provided. Cross-sectional weights apply to the entire file of all those who completed an interview, either community or facility. These cross-sectional weights can be used for making estimates of the population enrolled in Medicare for the entire calendar year.

To enable SUDAAN (Professional Software for SURvey DATa ANALysis for Multi-stage Sample Designs) users to compute population estimates and the associated variance estimates, two variables have been included in this record, SUDSTRAT AND SUDUNIT. Please see Section 6 for a further discussion about weights and estimation using these files.

Survey Longitudinal Weights Records (RICs X2, X3, and X4)

The first set of longitudinal weights (RIC X2) applies to sample persons who completed fall-round interviews in the current and the first preceding year. This set of weights can be used to study data trends over a two-year period. By applying these weights to data in the current and first preceding year, users will be able to estimate change among the Medicare population who were alive for the full two-year period.

The second set of longitudinal weights (RIC X3) applies to sample persons who completed fall-round interviews in the current and the two preceding years. This set of weights can be used to study data trends over a three-year period. By applying these weights to data in the current and the two preceding years, users will be able to estimate change among the Medicare population who were alive for the full three-year period.

The third set of longitudinal weights (RIC X4) applies to sample persons who completed fall-round interviews in the current and the three preceding years. This set of weights can be used to study data trends over a four-year period. By applying these weights to data in the current and the three preceding years, users will be able to estimate change among the Medicare population who were alive for the full four-year period.

To enable SUDAAN (Professional Software for SURvey DATA ANalysis for Multi-stage Sample Designs) users to compute population estimates and the associated variance estimates, two variables have been included in these records, SUDSTRAT and SUDUNIT. Please see Section 5 for a further discussion about weights and estimation using these files.

It should be noted that this file and these longitudinal weights are not appropriate for doing mortality studies, an area of considerable interest. This file does not include those who may have died during calendar year 2002 prior to the fall interview. Analysts interested in this topic are encouraged to use the annual Cost and Use files which target the ever-enrolled population.

Claims Records (DME, HHA, HSP, INP, OTP, PHY, SNF)

The following rules were used to select bill and claims records for this file.

- Inpatient bills were included if the discharge or "through" date fell on or after January 1, 2006 and on or before December 31, 2006.
- Skilled nursing facility bills were included if the admission or "from" date fell on or after January 1, 2006 and on or before December 31, 2006.
- Home health agency and outpatient facility bills were included if the "through" date fell on or after January 1, 2006 and on or before December 31, 2006.
- Hospice bills were included if the admission or "from" date fell on or after January 1, 2006 and on or before December 31, 2006.

Section 3: Notes on Using the Data

- Physician or supplier claims were included if the latest "service thru" date fell on or after January 1, 2006 and on or before December 31, 2006.
- Durable medical equipment (DME) claims were included if the latest "service thru" date fell on or after January 1, 2006 and on or before December 31, 2006.

A total of 2,884 (about 18 percent) of the sample people did not use Medicare reimbursed services in a fee-for-service setting in 2006; consequently, there are no bill records for them in this file. These individuals may have used no services at all, services only in a managed care plan, or services provided by a payer other than Medicare. For the other 12,886 individuals in the sample, we have captured bills meeting the date criteria, processed and made available by CMS through June 2007.

Medicare Current Beneficiary Survey

CY 2006 Access to Care

Section 4: Questionnaires

This section contains copies of the community and facility questionnaires that were administered during Round 46 of the Medicare Current Beneficiary Survey. Round 46 is the Fifteenth annual update of information on Medicare beneficiaries' access to care. The questionnaires are similar in content and sequence of events; however, they differ in how they are administered. Of special note, RIC KN and RIC PD data were collected in the 2007 January-April round and RIC RX data was collected in the 2007 May-August round.

Because the questionnaires are conducted using CAPI, the questionnaires actually exist only as a computer-program, and it is impossible to replicate it exactly in hard copy. The version represented here lists the questions, verbatim, and shows the skip patterns. It also displays instructions to the programmers (enclosed in boxes), to the program, and to the interviewer. Although these instructions would be hidden from the respondent, they have been retained in this copy because they are important for understanding the flow of the questionnaire and for establishing logical links between questions.

Questions in all of the questionnaires are preceded by a number, which is cross-referred to variables in the codebook (Section 2). Since more than one variable may be collected in response to one question, each question has also been annotated with all of the variable names associated with it. Hypertext links both the annotated variables in the questionnaire and the codebook.

Community Component

The community component is conducted in the home of the respondent.

Components of the Community Questionnaire

The community instrument consists of the following components:

- Initial interview questionnaire
- Core questionnaire
- Supplement to the core questionnaire
- Interviewer remarks questionnaire

Initial interview questionnaire

This baseline questionnaire is used for the first interview when a sample person is added to the survey, that is, Round 37 for the 2003 panel, Round 40 for the 2004 panel, Round 43 for the 2005

Section 4: Questionnaire

panel, Round 46 for the 2006 panel. In the initial interview, we collect information about the national origin, age, education and income of the sample person. The interviewer also verifies the sample person's address and telephone number and obtains the names and addresses of people who might be willing to serve as proxy respondents. The interviewer also uses this opportunity to acquaint the respondent with the intent of the survey and to familiarize him or her with the MCBS calendar, and to emphasize the importance of keeping accurate records of medical care and expenses.

In subsequent interviews, some of the information collected in the initial interview will need to be updated. For example, the sample person's designation of his or her race is not likely to change, and will not be asked about again. On the other hand, the sample person's address or telephone number may change, so this information is verified in every interview, and updated when necessary.

Core questionnaire (community)

NOTE: This release does not include any cost or utilization information from the core questionnaire.

The core questionnaire is the major component of the community instrument. The questions focus on the use of medical services and the resulting costs, and are asked in essentially the same way each and every time the sample person is interviewed (after the first time). In each interview, the sample person is asked about new encounters, and to complete any partial information that was collected in the last interview. For example, the sample person may mention a doctor visit during the "utilization" part of the interview. In the "cost" section, the interviewer will ask if the sample person has any receipts or statements from the visit. If the answer is "yes", the interviewer will record information about costs from the statements, but if the answer is "no," the question will be stored until the next interview.

In Round 46, only persons in the longitudinal sample (that is, 2003, 2004, and 2005 panels) were interviewed with this questionnaire.

Supplement to the core questionnaire (community)

Supplemental questions are added to the core questionnaire to gather information about specific topics. The Round 46 supplement focuses on health status and access to care. It includes questions about the sample persons' general health (including standard measures such as IADLs and ADLs), their sources of medical care, and their satisfaction with that care.

Interviewer remarks questionnaire

The interviewer completes this questionnaire after every interview with the sample person. The interviewer is asked to evaluate the sample person's ability to respond to the questionnaire and to provide some information about the interview (for example, if a proxy answered the questionnaire, the interviewer provides reasons why the proxy was necessary). The interviewer is also encouraged

to provide comments that will assist the interviewer in remembering unique facts about the sample person, such as hearing or vision impairments, or that the sample person cannot read.

Table 5.1 - Components of the Community Questionnaire

In addition to the community component questionnaires used during Round 46, this table displays the special supplemental questionnaires that were used in Rounds 47 and 48.

Certain procedural or operational management sections are collected specifically to manage the collection process. These sections do not directly cross over to the final data file. They include:

AH	Contact Information for HA Facilities
CL	Closing Materials
IR	Interviewer Remarks
UPD	Name/Address Update
CT	Interrupt
CE	Restart
EX	Exit

The Initial, Core, Supplement to the Core, and Special Supplemental questionnaires comprise the remaining sections. As discussed earlier in this document, the cost and utilization sections are not included in this data file. The files below are split into two columns. The first column represents those sections that are included in this data file. The second column represents those sections that are not included in this data file.

* Summary sections - Updates and corrections are collected through the summaries. The respondent is handed a hard copy summary of information gathered in previous interviews, and is asked to verify the material. Changes are recorded if the respondent notices information that is not accurate.

Included Sections

IN	Introduction
ENS*	Enumeration
EN	Enumeration
HIS*	Health Insurance Summary
HI	Health Insurance
AC	Provider Probes/Access to Care
HF	Health Status & Functioning
SC	Satisfaction with Care
US	Usual Source of Care
HA	Housing Characteristics

NE	New Enrollee Supplement
DI	Demographics/Income (Supplemental Sample Only)
KN	Knowledge & Needs Supplement
PD	Prescription Drug Supplement
RX	Drug Coverage Supplement

Excluded Sections

UTS*	Utilization Summary
DU	Dental Utilization and Events
ER	Emergency Room Utilization & Events
IP	Inpatient Hospital Utilization & Events
OP	Outpatient Hospital Utilization & Events
IU	Institutional Utilization
HHS*	Home Health Utilization Summary
HH	Home Health Utilization & Events
MP	Medical Provider Utilization & Events
OM	Other Medical Expenses Utilization
PMS*	Prescribed Medicine Summary
ST	Charge Questions (Statement Series)
PM	Prescribed Medicine Utilization
NS	Charge Questions (No Statement Series)
CPS*	Charge/Payment Summary

Facility Questionnaire

Interviews are conducted in the facility where the respondent is residing at the time of the interview. Information is obtained from facility records; therefore, the beneficiary is never interviewed directly. It was decided early in the design of the MCBS not to attempt interviews with sample persons in facilities, or with their family members. For that reason, the facility questionnaires do not ask about attitudes or other subjective items.

If an institutionalized person returns to the community, a community interview is conducted. If the sample person spent part of the reference period in the community and part in an institution, then a separate interview is conducted for each period of time. In this way, a beneficiary is followed in and out of facilities and a continuous record is maintained regardless of the location of the respondent.

The CY 2006 Access to Care release is intended to serve as a "snapshot" of the sample person at one point in time during Round 46. For this reason, we have selected the latest interview in the round to

represent the entire round whenever the sample person was encountered in more than one setting in Round 46.

Components of the Facility Questionnaire

The facility instrument consists of the following components:

- Facility eligibility screener
- Initial (baseline) questionnaire
- Core questionnaire
- Supplement to the core questionnaire

Facility eligibility screener

This questionnaire gathers information about the facility to determine the facility type. The initial interview is conducted with the facility administrator. All other interviews are conducted with the staff designated by the director. A facility screener is administered upon the sample person's admission to a new facility, and once a year thereafter (in Rounds 37, 40, 43, and 46) to capture any changes in the facility's size or composition. The screener is not administered if the sample person simply re-enters the same facility.

Initial (baseline) questionnaire (facility)

This questionnaire gathers information on the health status, insurance coverage, residence history and demographics of the sample person. This questionnaire is administered the first time the sample person is admitted to a facility.

Core questionnaire (facility)

This questionnaire parallels the core questionnaire for the community, collecting information about use of medical services and their associated costs, including the facility cost. Like its community counterpart, this questionnaire is administered in each and every interview after the first one, as long as the sample person continues to reside in the facility.

Supplement to the core questionnaire (facility)

This questionnaire is asked once a year (in Rounds 37, 40, 43 and 46) to update our information about the sample person's health status. It includes questions about the sample person's general health (including standard measures such as IADLs and ADLs), but excludes the questions about access and the subjective questions about satisfaction with care.

Table 5.2 - Components of the Facility Questionnaire

NOTE: This release contains information from only those sections marked with an arrow (→).

Facility Eligibility Screener

→ FQ Facility questions

Initial interview (facility)

→ A Demographic/Income
→ B Residence History
→ C Health Status and Functioning
→ D Health Insurance
L Tracing and Closing

Core questionnaire (facility)

→ A Residence History
B Provider Probes
C Medicine Summary
D Inpatient Hospital Stays
E Medical Charges
F Tracing and Closing

Supplement to the core (facility)

→ C Health Status and Functioning
→ D Health Insurance

Medicare Current Beneficiary Survey CY 2006 Access to Care

Section 5: Sample Design and Guidelines for Preparing Statistics

This section opens with a description of the population covered by the 2006 Access to Care release and a comparison of this “view” with others that are frequently used for analyzing the Medicare program. Next the targeted population is discussed in terms of the sampling strata. This is followed by a general discussion of the selection of the original and supplemental samples. Next appears a description of primary sampling units (PSU) and clusters of zip codes within PSU. Following is a general review of person level response rates, completed interviews by age strata, and selected item nonresponse rates. Guidelines for preparing population estimates using full sample weights and variance estimates using replicate weights are then reviewed.

Medicare population covered by the 2006 public use data

The calendar year 2006 MCBS public use data are focused on Medicare beneficiaries residing in the United States or Puerto Rico who were enrolled in one or both parts of the program throughout calendar year 2006. This “always enrolled” population includes individuals enrolled on January 1, 2006 who remained enrolled through the end of December. Excluded are the following categories of Medicare enrollees:

- 1) residents of foreign countries and U.S. possessions and territories other than Puerto Rico;
- 2) persons who became enrolled after January 1, 2006; and
- 3) persons who disenrolled or died prior to the end of December 2006.

NOTE: A small number of sample people included in this file (556) died during 2006 subsequent to their Round 46 interview and are a subset of group 3) above. A discussion of how to subset this file to get the “always enrolled” population is included in Section 3, “Notes on Using the Data”, under ‘Specific Variables – Key Record (RIC K)’.

While the “always enrolled” population differs from other views of the Medicare population commonly generated from CMS files or encountered in CMS publications such as “ever enrolled” or “mid-point enrollment”, the concept of “always enrolled” is consistent with the familiar concept of being exposed or “at risk” for using services for the entire 12-month period.

Table 5.1 shows data from CMS's 5-percent HISKEW file (health insurance skeleton write-off), which contains selected demographic and coverage information on a 5-percent sample of Medicare enrollees. Data for the targeted population are arrayed by age, sex, race using these three views: persons “ever-enrolled”, persons enrolled as of the “mid-point of the year” (July 1), and persons

Section 5: Sample Design and Estimation

Table 5.1 2006 Medicare population, by sex, race and age

Sex	Race	Age	Always Enrolled	July 1 Midpoint	Ever Enrolled
Total			40,674,620	42,970,600	45,092,280
Males	Black	0-44	182,060	193,680	204,740
		45-64	439,160	468,160	493,980
		65-69	350,200	392,880	427,320
		70-74	313,560	321,120	329,140
		75-79	208,980	216,420	223,660
		80-84	127,860	134,260	141,100
		85+	95,780	102,340	110,560
	Non-black	0-44	699,520	745,400	788,060
		45-64	2,110,960	2,255,660	2,384,340
		65-69	3,628,380	4,091,360	4,485,160
		70-74	3,474,860	3,532,840	3,589,260
		75-79	2,841,120	2,910,700	2,984,180
		80-84	1,981,480	2,061,300	2,144,880
		85+	1,431,460	1,543,160	1,659,700
Females	Black	0-44	148,880	161,260	171,620
		45-64	445,460	475,680	502,740
		65-69	444,500	497,660	539,460
		70-74	434,980	441,500	448,820
		75-79	336,780	344,440	352,180
		80-84	246,780	255,120	263,340
		85+	260,700	277,440	294,860
	Non-black	0-44	570,160	610,020	644,840
		45-64	1,846,760	1,966,120	2,072,100
		65-69	4,008,780	4,516,520	4,950,120
		70-74	4,107,040	4,156,180	4,202,020
		75-79	3,745,240	3,808,800	3,877,020
		80-84	3,071,220	3,158,040	3,250,480
		85+	3,121,960	3,332,540	3,556,600
Male total			17,885,380	18,969,280	19,966,080
Female total			22,789,240	24,001,320	25,126,200
Black total			4,035,680	4,281,960	4,503,520
Non-black total			36,638,940	38,688,640	40,588,760

Based on September 2007 HISKEW files, inflated to 100 percent. "Always Enrolled" data are estimated .

Section 5: Sample Design and Estimation

“always enrolled”. We have included these relationships to allow users to compare the population represented by this release to the more frequently used views of the Medicare population.

(It should be noted that the other series of files produced from the MCBS, known as “Cost and Use”, contains data composed to represent the ever-enrolled population to better capture total Medicare and other expenditures for a given year. A discussion of how the ever-enrolled population was composed for a given period to capture total utilization (covered and non-covered) and expenditures (Medicare and other) for that period is presented in the documentation for the 2005 file for that series.)

Targeted population and sampling strata

The targeted population for Round 1 of the MCBS consisted of persons enrolled in one or both parts of the Medicare program, that is, Part A (Hospital Insurance) or Part B (Supplementary Medical Insurance) as of January 1, 1991, and whose address on the Medicare files was in one of the 50 states, the District of Columbia, or Puerto Rico. Correspondingly, for Rounds 4, 7, 10, ... and 46 the targeted populations included those individuals enrolled as of January 1, 1992,¹ 1993,² 1994,³ ... and 2006 respectively.

Table 5.2 Targeted number of MCBS sample persons with complete annual utilization and expenditure data by sampling stratum

Age group	Target
Total	12,000
0 - 44	1,000
45 - 64	1,000
65 - 69	2,000
70 - 74	2,000
75 - 79	2,000
80 - 84	2,000
85 +	2,000

In general, the targeted universe for the MCBS is divided into seven analytic strata based on age as of December 31, 2006. The age categories are: 0 to 44, 45 to 64, 65 to 69, 70 to 74, 75 to 79, 80 to 84, and 85 or older. The goal of the sample design is to obtain complete annual data on health care use of both Medicare covered as well as non-covered services and the associated costs by source of

Section 5: Sample Design and Estimation

payment on 12,000 beneficiaries per year, with 2,000 for each of the elderly strata and 1,000 for each of the disabled strata. See Table 5.2.

(Note: Starting with the implementation of the rotating panel design described earlier in the Introduction Section, it was estimated that it would be necessary to interview roughly 16,000 sample persons in each of the fall rounds 1994 - 2006 (Rounds 10, 13, 16, ... and 46) in order to meet the targets for complete annual utilization and expenditure data for 1994 through 2006. The annual sample is distributed across four currently active panels of which the oldest will be replaced by a new panel in the following year. This use of such overlapping of panels will continue indefinitely under the rotating panel design for each fall interview session. See Table 5.3 for actual number of Round 46 completes by age stratum.)

Beneficiaries for the original sample (Round 1), the first supplement (Round 4), and the second supplement (Round 7) were selected from the standard 5-percent sample of CMS's Enrollment Data Base (EDB). The decision to select the MCBS sample from within the standard 5-percent CMS sample was based on considerations of convenience. The standard 5-percent sample has been used for many research projects involving the Medicare population, and data files have been constructed to allow access to the claims for this 5-percent sample.

The development of the National Claims History File makes the claims data generally available for the entire population, not just those individuals included in the standard 5-percent sample. In addition, for some PSU areas, the number of beneficiaries within the standard 5-percent sample is so small that the list of potential sample people can become exhausted over time. As a result, beginning in Round 10, the beneficiary samples have been drawn from 5-percent samples other than the standard 5-percent sample.

Primary and Second Stage Sampling Units

All of the continuing panels in the 2006 Access to Care release are spread across the redesigned sample of 107 primary sampling units (PSUs) selected in 2001. These PSUs are geographically dispersed and consist of metropolitan areas and groups of non-metropolitan counties. The 2001 panel, which was the first panel to be selected from the redesigned MCBS PSUs, is concentrated in 1,209 second-stage units consisting of clusters of ZIP code areas. ZIP code cluster augmentation for the last four years is as follows:

- ❖ The 2003 panel had nineteen new ZIP code clusters selected to augment those selected previously resulting in a total sample of 1,237 clusters.
- ❖ The 2004 panel was concentrated in the same 1,237 ZIP code clusters used to select the 2003 panel.

Section 5: Sample Design and Estimation

- ❖ For the 2005 panel, five new ZIP code clusters were selected, resulting in a total sample of 1,242 clusters.
- ❖ The 2006 panel had two new ZIP code clusters selected, resulting in a total sample of 1,244 clusters.

With the rotating panel design, the PSU redesign is transparent to data users and no special processing is required. For more details on the PSU redesign, see Lo, Chu, and Apodaca (2002).

All of the ZIP cluster samples were selected from CMS' master file of beneficiaries enrolled in Medicare, using the beneficiary's address recorded in that file as of March of the year the individual was selected to be in the sample.

Sample selection

The MCBS sample is designed to provide nearly self-weighting (i.e., equiprobability) samples of beneficiaries within the age strata. Within the selected PSUs and ZIP clusters, a systematic sampling scheme with random starts is employed. A brief summary of the 2003 – 2006 panels is given below.

In 2003 a supplemental sample of 6,000 beneficiaries was added to the sample for Round 37. The 2003 supplemental MCBS sample included newly enrolled beneficiaries (i.e., beneficiaries who were enrolled during the period February 2002 through January 2003) as well as previously enrolled beneficiaries (i.e., beneficiaries who were enrolled on or before January 2002). The 2003 Access to Care questions were administered in September through December 2003 as part of the Round 37 interview for the continuing sample, and as part of the initial interview for the 2003 supplemental sample.

In 2004 a supplemental sample of 6,342 beneficiaries was added to the sample for Round 40. The somewhat larger sample size used in 2004 was intended to partially offset expected sample losses due to projected higher rates of non-response. The 2004 supplemental MCBS sample included newly enrolled beneficiaries (i.e., beneficiaries who were enrolled during the period February 2003 through January 2004) as well as previously enrolled beneficiaries (i.e., beneficiaries who were enrolled on or before January 2003). The 2004 Access to Care questions were administered in September through December 2004 as part of the Round 40 interview for the continuing sample, and as part of the initial interview for the 2004 supplemental sample.

In 2005 a supplemental sample of 6,565 beneficiaries was added to the sample for Round 43. As in 2004, the somewhat larger sample size used in 2005 was intended to partially offset expected sample losses due to projected higher rates of non-response. The 2005 supplemental MCBS sample included newly enrolled beneficiaries (i.e., beneficiaries who were enrolled during the period February 2004 through January 2005) as well as previously enrolled beneficiaries (i.e., beneficiaries

Section 5: Sample Design and Estimation

who were enrolled on or before January 2004). The 2005 Access to Care questions were administered in September through December 2005 as part of the Round 43 interview for the continuing sample, and as part of the initial interview for the 2005 supplemental sample.

Finally, in 2006 a supplemental sample of 6,676 beneficiaries was added to the sample for Round 46. The somewhat larger sample size used in 2006 was intended to partially offset expected sample losses due to projected higher rates of non-response. The 2006 supplemental MCBS sample included newly enrolled beneficiaries (i.e., beneficiaries who were enrolled during the period February 2005 through January 2006) as well as previously enrolled beneficiaries (i.e., beneficiaries who were enrolled on or before January 2005). The 2006 Access to Care questions were administered in September through December 2006 as part of the Round 46 interview for the continuing sample, and as part of the initial interview for the 2006 supplemental sample.

Response Rates

By Round 46, 65 percent of the 2003 panel were still in a formal responding status (that is, either the SP was alive and still participating or had died after Round 37 but left behind a cooperative proxy for the collection of data on the last months of life) or had participated in the survey until death, leaving enough data to estimate the last months of life. For the 2004 and 2005 panels, the corresponding figures were 68 and 72 percent, respectively. The 2006 panel (Round 46) had an initial response rate of 83 percent.

There were 3,131 interviews successfully completed at Round 46 with still-living members of the 2003 panel. For brevity, we refer to these 3,131 interviews as “live completes”. For the 2004, 2005 and 2006 panels there were 3,468, 3,972, and 5,199 live Round 46 completes, respectively.

Completed interviews by sampling strata

Table 5.3 lists the number of completed interviews for Round 46 for the continuing (Rounds 37, 40, and 43) and supplemental (Round 46) panels by age strata (age as of July 1, 2006). Under the rotating panel design, the beneficiaries selected in Round 34 were released from the sample prior to Round 46.

Section 5: Sample Design and Estimation

Table 5.3 MCBS Round 46 completed interviews for the continuing and supplemental panels

Age group	- Round 46 Completed Interviews -		Total
	Continuing	Supplemental	
Total	10,571	5,199	15,770
0 – 44	961	500	1,461
45 – 64	915	446	1,361
65 – 69	1,790	1,173	2,963
70 – 74	1,772	742	2,514
75 – 79	1,776	831	2,607
80 – 84	1,752	853	2,605
85 +	1,605	654	2,259

Item Non-response

As in any other survey, some respondents could not, or would not, supply answers to some questions. Item non-response rates are generally low in the CY 2006 Access to Care release, but the analyst still needs to be aware of the missing data and be cautious about patterns of nonresponse.

Some of the missing data is attributable to the fact that some of the community interviews and all of the facility interviews are conducted through a proxy respondent. In other words, the respondent may not have had knowledge of the information sought on the sample person. In other situations the respondent may have simply refused to answer. While no effort was made to specifically impute missing data for the Access to Care variables, where possible, much of the data not collected in the current round has been filled in through editing to earlier files or through the use of files that have been imputed.

Each user can decide how to handle the missing data. One simple approach is to delete records with missing data. The cumulative effect of deleting each record with any missing data can significantly reduce the data available for analysis. Another approach is to create an “unknown” or “missing” category within each variable distribution. This approach retains more observations than the first approach. Neither of these approaches, however, compensates for potential biases due to non-response.

There are other better (though more complicated) alternatives for handling cases with missing data. One is to impute the missing data. This can be done fairly easily in such a way as to improve univariate tabulations, but techniques that retain correlation structure for multivariate analyses are extremely complex. For more discussion of imputation, the user is referred to Kalton and Kasprzyk

(1986). An alternative is model-based estimation where a joint mechanism is hypothesized that underlies both the substantive data and the missing data structure. For a discussion of this technique, see Little and Rubin (1987).

Preparing Statistics (Using the Full Sample Weights)

Four sets of final “full sample” Round 46 weights have been provided for different types of statistics. (The term “full sample” is used to distinguish these weights from the replicate weights discussed in the next section). One set of weights is labeled **CS1YRWGT** (RIC X). **CS1YRWGT** is a cross-sectional weight and applies to both the continuing sample and to the supplemental sample, which was fielded for the first time in Round 46. This weight is intended for use in cross-sectional statistics involving the total (combined) Round 46 sample. Each weight is greater than zero for all 15,770 beneficiaries (live completes) on the file. **CS1YRWGT** should be used to make estimates of the levels of access to care for the Medicare population alive in the fall of 2006.

The second set of weights is labeled **L4YRSWGT** (RIC X4). These weights are referred to as “three-year backward longitudinal” weights, and are intended for use in longitudinal analyses involving respondents in the 2003 panel who completed Round 37, 40, 43 and 46 interviews. The longitudinal weight, **L4YRSWGT**, does not apply to the 2004, 2005 and 2006 panels. This weight should only be used when the CY 2003 (Round 37), CY 2004 (Round 40), CY 2005 (Round 43) and CY 2006 (Round 46) Access to Care releases have been merged together at the beneficiary level. Only those beneficiaries who completed the Round 37 and Round 46 interviews have positive numeric values of **L4YRSWGT**. This weight can then be used to make estimates of changes in characteristics (say, health status) or attitudes (say, satisfaction with care) of a given subset of the population between the initial interview round (Round 37) and the final interview round (Round 46). There are 3,131 beneficiaries who completed the Round 37 and 46 interviews; these beneficiaries can be identified by selecting cases for which **D_SMPTYP**=’01’ (RIC K, Key Record). However, not all of the 3,131 beneficiaries with a valid **L4YRSWGT** completed all of the interviews between Rounds 37 and 46. As can be seen in Table 5-4, 32 beneficiaries did not complete the 2004 (Round 40) interview, 21 did not complete the 2005 (Round 43) interview, and two did not complete the 2004 and 2005 interviews. Analysts should be aware of this when using **L4YRSWGT**. For example, depending on the type of longitudinal analyses to be conducted, it may be desirable to compensate for the missing data by imputation or additional weighting adjustments.

The third set of weights is labeled **L3YRSWGT** (RIC X3). These weights are referred to as “two-year backward longitudinal” weights and are intended for use in longitudinal analyses involving respondents in the 2003 and 2004 panels who completed interviews in at least one prior round. The longitudinal weight, **L3YRSWGT**, does not apply to the 2005 or 2006 panel cases. This weight should only be used when the CY 2004 (Round 40), CY 2005 (Round 43) and CY 2006 (Round 46)

Section 5: Sample Design and Estimation

Access to Care releases have been merged together. Records must be merged at the beneficiary level. This weight can then be used to make estimates of year-to-year changes from 2004 to 2006. As indicated in Table 5-4, **L3YRSWGT** is greater than zero for 3,468 persons in the 2004 panel who completed the Round 40 and 46 interviews, and for 3,097 persons in the 2003 panel who completed the Round 40 and 46 interviews. However, 34 persons in the 2003 panel did not complete the Round 40 interview. **L3YRSWGT** is inapplicable (".") for the respondents in the 2005 and 2006 panels. There are 6,599 beneficiaries with valid **L3YRSWGT**; these beneficiaries are identified as **D_SMPTYP='01'** or **D_SMPTYP='02'** (RIC K, Key record).

The fourth set of weights is labeled **L2YRSWGT** (RIC X2). These weights are intended for use in longitudinal analyses involving respondents from the 2003, 2004 and 2005 panels who completed an interview in Round 46 and at least one other interview prior to Round 46. The longitudinal weight, **L2YRSWGT**, does not apply to the 2006 panel cases. This weight should only be used when the CY 2005 (Round 43) and CY 2006 (Round 46) Access to Care releases have been merged together at the beneficiary level. This weight can then be used to make estimates of year-to-year changes, such as the number of persons who went from being very satisfied with their care to being dissatisfied with their care. As indicated in Table 5-4, **L2YRSWGT** is greater than zero for 3,972 persons in the 2005 panel who completed the Round 43 and 46 interviews, 3,445 persons in the 2004 panel who completed the Round 43 and 46 interviews, and 3,108 persons in the 2003 panel who completed the 43 and 46 interviews however, a total of 46 beneficiaries in the 2003 and 2004 panels completed the Round 46 interview, but not the Round 43 interview. This weight; it is inapplicable (".") for the 2006 panel. There are 10,571 beneficiaries with a valid **L2YRSWGT**; these beneficiaries are identified as **D_SMPTYP='01'**, **D_SMPTYP='02'** or **D_SMPTYP='03'** (RIC K, Key record).

Although it is possible to create some cross-sectional estimates using **L4YRSWGT**, **L3YRSWGT**, or **L2YRSWGT** and create longitudinal estimates using **CS1YRWGT**, both of these actions are strongly discouraged. In general, estimates of the same population statistic produced using the two types of weights (that is, cross-sectional and longitudinal) will differ systematically. When **L4YRSWGT**, **L3YRSWGT**, or **L2YRSWGT** are used for cross-sectional estimation, recently enrolled beneficiaries will not be represented. **CS1YRWGT** cannot in general be used for longitudinal estimation since they include adjustments for multiple representation that are not appropriate for longitudinal analyses. .

Section 5: Sample Design and Estimation

Table 5-4. Number of cases with positive longitudinal weight in the 2006 access-to-care file, by type, availability of access-to-care data, and panel

Type of weight	Available access-to-care data				No. cases by panel			
	Round 37 (2003)	Round 40 (2004)	Round 43 (2005)	Round 46 (2006)	Total	2003	2004	2005
1-year backward longitudinal (L2YRSWGT)			√ —	√ √	10,525 46	3,108 23	3,445 23	3,972 0
2-year backward longitudinal (L3YRSWGT)		√	√	√	6,521	2,076	3,445	
		√	—	√	44	21	23	
		—	√	√	32	32	0	
		—	—	√	2	2	0	
3-year backward longitudinal (L4YRSWGT)	√	√	√	√	3,076	3,076		
	√	√	—	√	21	21		
	√	—	√	√	32	32		
	√	—	—	√	2	2		

√ Indicates that survey data are available in MCBS access-to-care data releases.

Variance Estimation (Using the Replicate Weights)

In many statistical packages, including SAS, the procedures for calculating sampling errors (e.g., variances, standard errors, etc.) assume that the data were collected in a simple random sample. Procedures of this type are not appropriate for calculating the sampling errors of statistics based upon a stratified, unequal-probability, multi-stage sample such as the MCBS. Thus, to permit the calculation of sampling errors, a series of replicate weights are included in the 2006 Access to Care release. Unless the complex nature of the MCBS is taken into account, estimates of the variance of a survey statistic may be biased downward.

The replicate weights included in the MCBS data files can be used to calculate standard errors of the sample-based estimates as described below. Just as there are four full sample weights for Round 46, one for cross-sectional analyses and three for longitudinal analyses, there are four corresponding sets of replicate weights.

The replicate cross-sectional weights are labeled **CS1YR001** through **CS1YR100** and may be found in the Cross-sectional Weights record (RIC X). The first set of replicate longitudinal weights is labeled **L4YRS001** through **L4YRS100** and may be found on the first Longitudinal Weights Record

(RIC X4). The second set of replicate longitudinal weights is labeled **L3YRS001** through **L3YRS100** and may be found on the second Longitudinal Weights Record (RIC X3). The third set of replicate longitudinal weights is labeled **L2YRS001** through **L2YRS100** and may be found on the third Longitudinal Weights Record (RIC X2).

Most commercial software packages today include techniques to accommodate the complex design, either through Taylor-expansion type approaches or replicate weight approaches. In addition, Westat maintains a package called WesVar. A free downloadable version can be obtained from Westat's home page: **WWW.WESTAT.COM**. Technical questions regarding the use of the MCBS replicate weights for variance estimation may be directed to David Ferraro at Westat, telephone 301.251.4261.

Another option is for the user to write a small custom program using a very simple algorithm. Let X_0 be an estimate of a statistic of interest formed using one of the full sample weights. Let X_1 through X_{100} be estimates (calculated by the user) of the same statistic of interest formed using the corresponding 100 replicate weights. The estimated variance of X_0 is then simply:

$$Var(X_0) = \frac{2.04}{100} \sum_{i=1}^{100} (X_i - X_0)^2$$

Consistency with Medicare Program Statistics and with Prior Access to Care releases

In general, MCBS estimates may differ from Medicare program statistics such as tabulations of the HISKEW. There are several reasons for the differences. The most important reason is that the EDB, and hence the HISKEW, includes people who are no longer alive. This may occur where people have entitlement, say for Part A only, and receive no Social Security check. When field-staff try to locate these beneficiaries for interviews, they establish the fact of these deaths. This over-reporting on the HISKEW files is under constant scrutiny, with efforts to modify CMS edit procedures that reject records from SSA that have valid dates of death but other erroneous information. Unrecorded deaths are still present on the EDB.

For cross-sectional estimates of the CY 1992 and CY 1993 Access to Care, special weighting procedures were used to force some MCBS estimates to agree with HISKEW tabulations. This was not the case for the CY 1991 release nor is it the case for CY 1994 through CY 2006.

Although revised 1992 weights have not been published, internal analyses indicate that the estimated mean reimbursement per enrollee in 1992 was slightly too low in all but the youngest age category. The estimate of mean total Medicare reimbursement amount in 1992 increased 0.25 percent from \$5,022 per beneficiary before reweighting to \$5,035 per beneficiary after reweighting. This downward bias was due to the presence of a relatively small number of individuals on the enrollment

Section 5: Sample Design and Estimation

database who are listed as alive when in fact they are dead (primarily individuals not receiving social security benefits but deemed “insured” for purposes of Medicare at the beginning of the program). Consequently, the number of enrollees was estimated to be too large. This effect was concentrated in the 85+ age category since that is where most of the unrecorded deaths were detected during survey field operations. Results of internal analysis indicate that the estimated proportion of beneficiaries in the 85 + age group decreased from 9.6 percent to 9.0 percent after reweighting.

Also, research has shown that MCBS estimates by Census Division (the nation is divided into nine of these areas) are subject to extremely high variance. Users are urged to collapse down to the standard four Census Regions for geographic analyses.

There was an error in the 1992 weighting that skewed sample estimates toward non-metropolitan characteristics. Results from the reweighting indicate that mean total 1992 Medicare reimbursement increased from \$4,238 to \$4,269 for non-metropolitan beneficiaries, but decreased from \$5,334 to \$5,303 for beneficiaries in metropolitan areas.

Finally, an error in sampling was discovered early in 1996 that resulted in the selection of 320 beneficiaries from ZIP clusters that are not in sampled PSUs. Of these missampled cases, 64 were sampled with the 1992 supplementary panel, 38 were sampled with the 1993 supplementary panel, 115 were sampled with the 1994 supplementary panel, and 103 were sampled with the 1995 supplementary panel. Most of these cases are located in Central Florida (92 cases) and Puerto Rico (181 cases). Missampled cases in the 1991-1993 panels were released at Round 16. Seven missampled cases in the 1992 panel remained in the sample at Round 16 but were released at Round 19 when this panel rotated out of the sample.

Weighting procedures for the 1996 Access to Care releases included a special adjustment to correct for the missampling. Weights for the previous Access to Care releases have not been revised to reflect this correction. Although preliminary analysis indicates that the effect on overall estimates is small, estimates for Hispanics appear to be more affected by the sampling errors, and analysts are advised to use caution when interpreting results for this subgroup that are obtained from earlier Access to Care releases.

References

- Kalton, G. and Kasprzyk, D. (1986), "The Treatment of Missing Survey Data," *Survey Methodology*, 12, 1-16.
- Little, R.J.A. and Rubin, D.B. (1987), *Statistical Analysis with Missing Data*, New York: John Wiley and Sons.
- Lo, A., Chu, A. and Apodaca, R. (2002). Redesign of the Medicare Current Beneficiary Survey Sample. *Proceedings of the Survey Research Methods Section of the American Statistical Association* [CD-ROM], Alexandria, VA: American Statistical Association.

Notes

1. Because people often are listed in the enrollment database before eligibility for Medicare benefits actually begins, the Round 4 supplement includes some beneficiaries who became eligible in February 1992 or later.
2. The primary goal of the survey is the dissemination of detailed calendar year utilization and expenditure data. Therefore, it was decided to delay sampling the January 1993 accretes to coincide with sampling for the balance of 1993 accretes. Since the target for the CY 1993 Access to Care file was all those enrolled for all of 1993, which includes the January 1993 accretes, a weighting adjustment was made so that the cross-sectional weights on that file do reflect the inclusion of this group.
3. The primary goal of the survey is the dissemination of detailed calendar year utilization and expenditure data. Therefore, it was decided to delay sampling the January 1994 accretes to coincide with sampling for the balance of 1994 accretes. Since the target for the CY 1994 Access to Care file was all those enrolled for all of 1994, which includes the January 1994 accretes, a weighting adjustment was made so that the cross-sectional weights on that file do reflect the inclusion of this group.