

Medicare Current Beneficiary Survey
CY 1994 Access to Care

Public Use File Documentation

Introduction

The accompanying public use file is the fourth 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 September through December of 1994 (or earlier rounds for some variables for individuals in the continuing sample), and augmented with Medicare claims and administrative data for calendar year (CY) 1994.

Purposes of the Survey

The MCBS is a continuous, multi-purpose 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). It is designed to support both the agency's and that office's functions. CMS's primary mission is administering the Medicare program (Health Insurance for the Aged and Disabled) and assisting the states in administering the Medicaid program (Grants to States for Medical Assistance Programs).

The Office of the Actuary is responsible for the following activities: promulgating the premiums and deductible and coinsurance amounts associated with coverage or use of covered services; estimating the cost of covered benefits for Medicare and Medicaid for the budgetary exercise; estimating the cost of noncovered services in response to legislative initiatives; projecting the long term financial soundness of the health insurance trust funds; estimating the cost of health reform; and estimating current and projected national and personal health care expenditures (the national health accounts).

Analysis of the accounts shows that in recent years there has been some slowing in the growth in health care expenditures. The decade from 1980 to 1990 experienced an average annual rate in the growth of national health expenditures of almost 11 percent. The average annual rate has slowed from 1990 to 1994 to 8 percent. However, health care spending continues to outpace the overall economy (as measured by the gross domestic product), which grow at an average rate of about 5 percent during the same time.

In addition, the share of the nation's total health care bill funded by the Federal, State and local governments through the Medicare and the Medicaid programs rose rapidly during the early part of this decade, from almost 27 percent in 1990 to 31 percent in 1994.

These trends, along with CMS's concerns about the quality of care beneficiaries are receiving, support the need for the collection of a wide variety of health related data. The MCBS is designed to aid in CMS's administration, monitoring, and evaluation of the Medicare and Medicaid programs. MCBS data will enable 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 prescription drugs and 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

Field work 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 over several 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 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 over-sampled to

permit more detailed analysis of these subpopulations. The sample size was designed to yield about 12,000 completed cases annually.

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 initially 1,163 geographic clusters (later expanded to 1,344 in Round 4, 1,399 in Round 7, and 1,432 in Round 10) randomly drawn within those PSUs. The sample was annually supplemented during the September through December interview periods (that is, Round 4, Round 7, and Round 10) to account for attrition (deaths, dis-enrollments, refusals, etc.) and newly enrolled persons.

The set of beneficiaries reported on in the 1994 Access Care to File consists of a random cross-section of all beneficiaries who were enrolled for one or both parts of the Medicare program as of January 1, 1994 and were alive and enrolled at the time of interview during the 1994 fall round (September - December). Their names were drawn through the use of a fairly complex selection algorithm that involves four different panels.

The initial large panel of 15,411 beneficiaries was fielded in the fall of 1991. Smaller supplemental panels were added in the fall of 1992 and 1993. These supplementary panels afforded a chance of selection to beneficiaries who became entitled to either part A or Part B benefits during 1991 and 1992 in addition to maintaining adequate sample sizes in the face of death and sample attrition. At the time that the first panel was fielded, no definite decision had been made on how many years to interview sample beneficiaries.

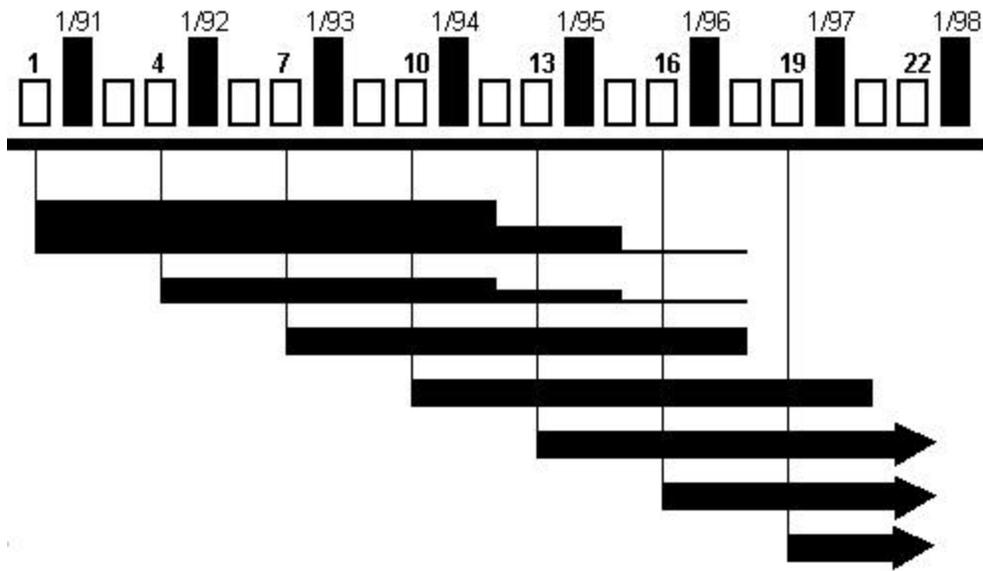
In 1993, a decision was made phase out the 1991, 1992 and 1993 Panels after no more than 6 years of interviews and to limit future panels to four years of interviews. This meant that the new sample to be selected for 1994 had to be designed like the 1991 Panel so that it could eventually replace it, rather than being narrowly focused as the 1992 and 1993 Panels were.

At the same time, a decision was made to increase the overall sample size in terms of interviews per year in order to allow the simultaneous interviewing of 4 panels, each starting with about 6400 sample beneficiaries. Accordingly, the 1994 Panel, introduced in the fall of 1994, consisted of 6390 beneficiaries, including a sample of those who became entitled during 1993.

In Round 10 (September-December 1994), we began implementation of the rotating panel process with the 1994 Panel. This group consisted of 6,390 beneficiaries, including a sample of those who became entitled during 1993 and on January 1, 1994. Once the rotating panel design is fully implemented, a panel will be followed for 12 interviews. There will be four panels active at any one time, and each panel will have approximately 4,000 active sample persons. New panels will be introduced each year in the fall round and the panel being replaced will be retired the following summer.

Because of the overlap between the new panel and the retiring panel, the number of interviews we conduct in a round increased from 12,000 to 16,000 starting in Round 10. Figure I.1, while not drawn to scale, gives a visual display of a timeline and shows the transition from a longitudinal to a panel design. The overlap during the transition and the simultaneous fielding of four panels during the fall round once the transition is implemented, is represented there.

Figure I.1 Transition to a Rotating Panel Design



The retiring panel (about 4,000 individuals) had abbreviated questionnaires administered beginning in Round 11 and, if necessary, Round 12 to complete the collection of medical events occurring in 1994. These sample persons were then rotated out of the study. These individuals will have participated a total 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).

Each fall, under the rotating panel design, a new panel will be introduced and each summer a panel will be retired. Thus, the sample will again be supplemented in Round 13 and approximately 4,000 of the ongoing sample will be retired from the study by Round 15. In Round 16 the remaining 4,000 will start the process of being rotated out. By Round 18 (the second round in 1997), the rotating panel design will be fully implemented; all of the original and first two supplements will have completed the study. This rotating panel sample design allows for both the eventual termination of participation in the study for individuals and for the completion of about 12,000 interviews for an ongoing study population.

As mentioned above, the phase-out of 4,000 individuals from the 1991 and 1992 panels was not scheduled until Round 12, while the first interview for the 1994 panel was conducted in Round 10. Therefore, the sample size at Round 10 was much larger than it was for previous fall rounds because of the overlapping panels. Table I.1 shows the number of respondents by age as of July 1, 1994. Differential sampling rates were used to obtain such large samples of the disabled and the oldest old.

 Table I.1 Number of completes at Round 10 by age category

Age as of 7/1/94	Number of completes at Round 10
Total	16,119
Under 45	1,287
45 to 64	1,477
65 to 69	3,004
70 to 74	2,594
75 to 79	2,703
80 to 84	2,649
85 and older	2,405

Complex Sample Design. In order to maximize the precision of statistics, the sample was concentrated in about 1500 ZIP code clusters within 107 Metropolitan Statistical Areas and clusters of nonmetropolitan 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 differential sampling 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 both the clustering and for the differential sampling and will give seriously incorrect standard errors, confidence intervals and p-values. Two packages that will give correct estimates of precision and significance are SUDAAN and WESVAR. See Section 6 for information on how to use these packages.

A concerted effort was made to keep response rates high at each round, knowing that cumulative attrition can become very serious in a panel survey. By Round 10, 69 percent of the survivors in the 1991 Panel were still responding. (The other 31 percent dropped out of the survey after providing anywhere from 0 to 9 interviews.) The comparable figures for the 1992, 1993 and 1994 Panels were 71%, 73%, and 83%. (Their rates are higher since they have not been asked yet for as many interviews.) The average response rate across the four panels at Round 10 was 73 percent.

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 10, 13 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.

The institutional questionnaires also differ from those used in the community because the former are administered using a conventional printed instrument with handwritten entries.

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 data base 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 Public Use File

The Access to Care Public Use 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 by-passes the reconciliation of the utilization and expenditure data collected in the survey with Medicare claims data.

The content of the Access to Care Public Use 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: A list of anomalies that exist in the data which were intentionally left as reported by the respondent ("No-Fix" edits).

Section 5: Hard copy versions of the questionnaires used in Round 10. The questionnaires have been annotated with variable names to associate the questions with the codebook.

Section 6: 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, unpublished 1994 data from the national health accounts.

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 after the targeted last accretion date, specifically between February and August 1992.

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 reflect the inclusion of this group.

Medicare Current Beneficiary Survey

CY 1994 Access to Care

Public Use File Structure

File specifications

The MCBS Calendar Year 1994 Access to Care PUF consists of a series of 36 separate datasets or files. Twenty-one of these datasets contain data on the MCBS sample persons; these files are the data files. The other 15 datasets contain SAS® code (SAS input statements, formats and labels) to facilitate the use of the data files by users who have access to a SAS mainframe environment. These are the README files.

Figure 1.1 shows file specifications such as file name, record counts, and the associated README file names.

Summary of the Data

The 21 data files represent completed Round 10 interviews with a sample of 16,119 Medicare beneficiaries, and supplemental information from CMS's Medicare files. Of these cases, 14,742 beneficiaries had community interviews and 1,377 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 14 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.

Section 1: File Structure

-Figure 1.1: File organization

File name	Record Counts
MCBS.readme.rick	
MCBS.readme.rica	
MCBS.readme.ric1	
MCBS.readme.ric2	
MCBS.readme.ric3	
MCBS.readme.ric4	
MCBS.readme.ric5	
MCBS.readme.ric6	
MCBS.readme.ric7	
MCBS.readme.ric8	
MCBS.readme.ricx	
MCBS.readme.ricx4	
MCBS.readme.ricx3	
MCBS.readme.ricx2	
MCBS.rick	16,119
MCBS.rica	16,119
MCBS.ric1	16,119
MCBS.ric2	16,109
MCBS.ric3	14,732
MCBS.ric4	16,119
MCBS.ric5	14,742
MCBS.ric6	1,377
MCBS.ric7	1,377
MCBS.ric8	16,119
MCBS.ricx	16,119
MCBS.ricx4	16,119
MCBS.ricx3	16,119
MCBS.ricx2	16,119
MCBS.readme.billrec	
MCBS.billrec.inp	5,190
MCBS.billrec.snf	1,042
MCBS.billrec.hsp	300
MCBS.billrec.hha	7,888
MCBS.billrec.otp	39,640
MCBS.billrec.phy	500,243
MCBS.billrec.dme	27,930

There are seven data files containing Medicare bill data. The MCBS.README.BILLREC file contains SAS input statements and labels (but no formats) for all seven bill record files.

As an illustration of the structure of the README files, Figure 1.2 is a copy of the README file for the Survey Enumeration record, RIC5.

Section 1: File Structure

Figure 1.2: Text of a Typical README file
(MCBS.README.RIC5 Illustrated)

```
INPUT @1 RIC $1.
      @2 FILEYR $2.
      @4 BASEID $8.
      @12 D_HHTOT 2.
      @14 D_HHREL 2.
      @16 D_HHUNRL 2.
      @18 D_HHCOMP 2.
      @20 D_HHLT50 2.
      @22 D_HHGE50 2.;;

PROC FORMAT;
VALUE HHCDFMT . = 'INAPPLICABLE'
      -8 = 'DONT KNOW'
      1 = 'NO ONE'
      2 = 'SPOUSE ONLY'
      3 = 'SPOUSE & OTHERS'
      4 = 'CHILDREN ONLY'
      5 = 'CHILDREN & OTHERS'
      6 = 'OTHERS ONLY'
      7 = 'NON RELATIVE';

VALUE PEOPLE 0 = 'NO ONE'
      1 = 'ONE PERSON'
      2 = 'TWO PEOPLE'
      .
      .
      .
      22 = 'TWENTY TWO PEOPLE';
COMMENT USE THIS TO SET LABELS ON THIS FILE;

LABEL
      RIC = 'RIC CODE FOR SURVEY ENUMERATION CODE'
      FILEYR = 'YY REFERENCE YEAR OF RECORD'
      BASEID = 'UNIQUE IDENTIFICATION NUMBER'
      D_HHTOT = 'TOTAL NUMBER OF PEOPLE IN HH'
      D_HHREL = 'NO. IN HH RELATED TO SP (INCLUDING SP)'
      D_HHUNRL = 'TOTAL NO. PEOPLE IN HH UNRELATED TO SP'
      D_HHCOMP = 'HOUSEHOLD COMPOSITION CODE'
      D_HHLT50 = 'NUMBER IN HH UNDER 50 (MAY INCLUDE SP)'
      D_HHGE50 = 'NO. IN HH 50 AND OVER (MAY INCLUDE SP)';
FORMAT D_HHTOT PEOPLE.
      D_HHREL PEOPLE.
      D_HHUNRL PEOPLE.
      D_HHCOMP HHCDFMT.
      D_HHLT50 PEOPLE.
      D_HHGE50 PEOPLE.;
```

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 14 data files in the survey and administrative summary data group:

- Key
- Administrative Identification
- Survey Identification
- Survey Health Status and Functioning
- Survey Access to Care
- Survey Health Insurance
- Survey Enumeration
- Survey Facility Residence History
- Survey Facility Identification
- Survey Interview
- Survey Cross-Sectional Weights
- Survey Longitudinal Weights (for individuals from the original sample who completed Round 1, Round 4, Round 7, and Round 10 interviews)
- Survey Longitudinal Weights (for individuals from the original or Round 4 supplemental samples who completed Round 4, Round 7 and Round 10 interviews)
- Survey Longitudinal Weights (for individuals from the original, Round 4, or Round 7 supplemental samples who completed Round 7 and Round 10 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
- Durable medical equipment

The bill records represent services provided during calendar year 1994 and processed by CMS in conjunction with our administrative functions. To facilitate analysis,

Section 1: File Structure

the Administrative Identification record contains a summary of the utilization that these bills present in detail.

All MCBS public use records begin with the same three variables: a record identification code (RIC), the version of the RIC (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 10, 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; 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 - Survey Health Status and Functioning	1 per respondent	1 per respondent
RIC 3 - Survey Access to Care	1 per respondent	none
RIC 4 - Survey Health Insurance	1 per respondent	1 per respondent
RIC 5 - Survey Enumeration	1 per respondent	none
RIC 6 - Survey Facility Residence History	none	1 per respondent
RIC 7 - Survey Facility Identification	none	1 per respondent
RIC 8 - Survey Interview	1 per respondent	1 per respondent
RIC X - Survey Cross-sectional Weights	1 per respondent	1 per respondent
RIC X4 - Survey R1/4/7/10 Longitudinal Weights	1 per respondent	1 per respondent
RIC X3 - Survey R4/7/10 Longitudinal Weights	1 per respondent	1 per respondent
RIC X2 - Survey R7/10 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 (eff. 10/93) *	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

Survey and Administrative Summary Data Files

File: KEY

RIC: K

Number of Records: 16,119 - 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 check-digit.

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: 16,119 - 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 Health Care Financing Administration. 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 1994.

Table 1.A - File Overviews

Survey and Administrative Summary Data Files

File: SURVEY IDENTIFICATION

RIC: 1

Number of records: 16,119 - 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.

File: SURVEY HEALTH STATUS AND FUNCTIONING

RIC: 2

Number of Records: 16,109 - 1 for each person who completed an interview (excludes the following BASEIDs: 00005845, 00021801, 00036263, 00044597, 00048612, 00075404, 00123510, 00153105, 00182763, 00192227)

Description: The SURVEY HEALTH STATUS AND FUNCTIONING file contains information 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 such as smoking, 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

Survey and Administrative Summary Data Files

File: SURVEY ACCESS TO CARE

RIC: 3

Number of Records: 14,732 - 1 for each community person who completed a interview (excludes the following BASEIDs: 00005845, 00021801, 00036263, 00044597, 00048612, 00075404, 00123510, 00153105, 00182763, 00192227)

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 1994 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: 16,119 - 1 for each person who completed an interview

Description: The SURVEY HEALTH INSURANCE file summarizes the health insurance information provided by the sample people.

NOTE: One derived variable, the summary insurance indicator, D_SUMINS, indicates the variety and number of policies reported by the sample person. Medicaid coverage and details of other types of coverage are also included. To limit the size of the RIC 4 record, only 5 private health insurance policies are detailed. Eleven individuals in the sample had more than 5 private health insurance policies. For those eleven people the total in the summary indicator is correct, but the number of plans detailed is less than the total.

Table 1.A - File Overviews

Survey and Administrative Summary Data Files

File: SURVEY ENUMERATION

RIC: 5

Number of Records: 14,742 - 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: 1,377 - 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.

File: SURVEY FACILITY IDENTIFICATION

RIC: 7

Number of Records: 1,377 - 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.

Table 1.A - File Overviews

Survey and Administrative Summary Data Files

File: SURVEY INTERVIEW

RIC: 8

Number of Records: 16,119 - 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 CROSS-SECTIONAL WEIGHTS

RIC: X

Number of Records: 16,119 - 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: SURVEY R1/4/7/10 LONGITUDINAL WEIGHTS

RIC: X4

Number of Records: 16,119 - 1 non-zero weight for each individual from the original sample who completed Round 1, Round 4, Round 7, and Round 10 interviews (8,237); missing value (".") for all others (7,882).

Description: The LONGITUDINAL WEIGHTS file provides longitudinal weights, including general purpose weights and a series of replicate weights.

Table 1.A - File Overviews

Survey and Administrative Summary Data Files

File: SURVEY R4/7/10 LONGITUDINAL WEIGHTS

RIC: X3

Number of Records: 16,119 - 1 non-zero weight for each individual from the original or Round 4 sample who completed Round 4, Round 7, and Round 10 interviews (9,773); missing value (".") for all others (6,346).

Description: The LONGITUDINAL WEIGHTS file provides longitudinal weights, including general purpose weights and a series of replicate weights.

File: SURVEY R7/10 LONGITUDINAL WEIGHTS

RIC: X2

Number of Records: 16,119 - 1 non-zero weight for each individual from the original, Round 4, or Round 7 sample who completed Round 7, and Round 10 interviews (11,371); missing value (".") for all others (4,748).

Description: The LONGITUDINAL WEIGHTS file provides longitudinal weights, including general purpose weights and a series of replicate weights.

Table 1.B - File Overviews
Medicare Utilization Data Files

File: HOSPITAL BILL

RIC: INP

Number of Records: 5,190

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: 1,042

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: 300

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.B - File Overviews

Medicare Utilization Data Files

File: HOME HEALTH BILL

RIC: HHA

Number of Records: 7,888

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: 39,640

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: 500,243

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.

Table 1.B - File Overviews
Medicare Utilization Data Files

-

File: DURABLE MEDICAL EQUIPMENT BILL

RIC: DME

Number of Records: 27,930

Description: Medicare DME Part B claims for the MCBS population. These records reflect claims for DME rentals and purchases which were processed beginning 10/93.

Medicare Current Beneficiary Survey
CY 1994 Access to Care

Codebook

This public use 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's National Claims History (NCH) database.

The first part of this section includes frequency tables 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 tables

The following tables list the variables in each of the records, give their physical location in the record, list their possible values and relate them to the questionnaires or to source CMS files.

The first part of the Medicare Current Beneficiary Survey public use file (that is, the survey and CMS summary data) is made up of 14 different types of records. The name of the record type being described is identified by name on the second line. The RIC or record identification code with the record type being described is shown on the third line on the right of the page under the page number. This will enable more rapid access to particular parts of the codebook.

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 SAS variable names. All variables that are preceded by the character "D_", such as D_SMPTYP are derived variables. The variables did not come directly from the survey data, but compiled from several survey variables. Variables preceded by the characters "H_" 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.

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

This column will be blank for variables that relate to neither the questionnaire nor 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.

Ty (Type) - This column identifies the type of variable, that is, 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: "." 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 HS16), he would not be asked if he smokes now (question HS17).

The codebook for the various survey and summary RICs is followed by a Variable Name Index that lists sequentially all variables in the codebook, source of information, pertinent RIC, and page within the codebook.

Table 2.1: Abbreviations Used to Identify Sections of the Questionnaires

Community Questionnaire

IN	Introduction
EN	Enumeration
HI	Health Insurance
AC	Access to Care
HS	Health Status and Functioning
SC	Satisfaction with Care
US	Usual Source of Care
DI	Demographics/Income
CL	Closing

Facility Questionnaire (Screener)

FQ

Facility Baseline Questionnaire

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

BILL DETAIL SEGMENT

Using the tables

The tables in the bill detail section describe the Medicare utilization files included on the public use 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 public use 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 public use file bill records and CMS's 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 PROCNT.

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 six characters, interpreted as YMMDD (two-digit year, 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.

Medicare Current Beneficiary Survey

CY 1994 Access to Care

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 and how field procedures affect the data collection process. We have included this section to address those questions.

This section is a collection of information about various data fields present in this public use 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

Various negative values are used to indicate missing data. For instance, for survey collected data, a value of -1 indicates that the variable is inapplicable. A variable is generally inapplicable because the question is not appropriate, for example, a question about hysterectomy when the respondent is a male. In this file, the value -1 has been replaced with SAS® standard missing values (" " or blank for character and "." for numeric). Other missing value codes used in the survey (-7 for "refused," -8 for "don't know," and -9 for "not ascertained") were not changed.

Dates

The CMS-derived dates of birth include century indicators and are in the format YYYYMMDD (4-digit year, 2-digit month and 2-digit day). Other dates in this public use release have been written as six numeric characters in the following form: YYMMDD (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 "9206-8" ("92" for the year, "06" for June, 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 public use release does not contain narratives. Instead, we have supplied codes that summarize the answer. Often there will be more than one code because the answer included several specific topics.

Specific variables - Key Record (RIC K)

There are 16,119 key records, one for each individual who completed an interview (14,742 community interviews **INTERVU="C"** and 1,377 facility interviews **INTERVU="F"**).

The facility interview was conducted whenever the 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 extend the focus of their studies of facility care by using information from the Survey Facility Identification Record. This record is present for each sample person for whom a facility questionnaire was administered.

Some sample people had more than one interview in this round. This release is a mix of the original 1991 sample (**D_SMPTYP="91"**), people who joined the survey in Round 4 (**D_SMPTYP="92"**), people who joined the survey in Round 7 (**D_SMPTYP="93"**), and people who joined the survey in Round 10 (**D_SMPTYP="94"**). Interviews are conducted for the continuing sample (that is, the original 1991, Round 4, and Round 7 samples) in as many settings as necessary, to create a seamless view of the entire round. Only one interview--facility or community--is conducted with the new (Round 10) sample people, depending on the situation in which we locate them. In the total group of 16,119 sample persons with a Round 10 interview, the great majority had one, and only one, interview.

To avoid duplication of data, the records in this file represent only the last interview in Round 10: 14,742 community and 1,377 facility interviews.

This record contains a special-purpose variable, **SURVIVE**. The Round 10 MCBS data are not suited for making estimates of the "ever enrolled" 1994 Medicare population because the sample does not include beneficiaries new to Medicare in 1994 nor those who died during the year prior to the fall round. For the same reason, point-in-time estimates are also inappropriate. However, an adjustment has been done to account for persons newly enrolled on January 1, 1994. It is appropriate, therefore, to use the Round 10 data to approximate the "always enrolled" 1994 Medicare population, that is, those beneficiaries who were alive and enrolled on or before January 1, 1994 and were still alive and enrolled on January 1, 1995. We estimate this population at 34.6 million beneficiaries. This group can be separated from the entire group of Round 10 interviews by selecting only beneficiaries who were enrolled before 1994 and survived until 1995 (**SURVIVE="Y"**).

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, 1994.

Four variables relating to the sample person's age are provided. Date of birth as reported by the respondent during the initial interview is recorded in the RIC 1 - Survey Identification record (**D_DOB**). Date of birth from the Medicare - Social Security Administration records is recorded in the Administrative Identification Record (**H_DOB**). The variable **H_AGE** represents the sample person's age as of December 31, 1994. The variable **H_STRAT** groups the sample persons by **H_AGE**. The variables **H_DOB**, **H_AGE**, and **H_STRAT** appear in the Administrative Identification record.

As of July 1, 1994, approximately 4.6 million enrollees or almost 13 percent of the Medicare population had their Part B and/or Part A premiums paid by a State agency (for the entire year an estimated 5.2 million persons ever-enrolled had their premium paid for at least one month). This process, called State buy-in, is tracked by CMS and is used as a general proxy for Medicaid participation. The variables that describe this participation (**H_MCSW** and **H_MCDE01 - H_MCDE12**) were derived through a match with CMS's enrollment database. The variable **H_MCSW** can be used when only an indication that the enrollee was a "buy-in" at some time during 1994 is needed for analysis. The monthly variables **H_MCDE01 - H_MCDE12** can be used for analyzing Medicaid eligibility at specific points in time.

Membership in Medicare managed care plans has been increasing faster than the overall Medicare population. During 1994 managed care enrollees increased from 2.7 million in January to 3.1 million in December. As of the mid-point of the calendar year, July 1, approximately 2.9 million or 8 percent of the Medicare population received Medicare benefits through coordinated care organizations such as an HMO which contracts directly with CMS to provide those services. Some of the beneficiaries in the MCBS sample belong to such organizations. The variables that describe this membership (**H_GHPSW** and **H_PLTP01 - H_PLTP12**) were derived through a match with CMS's 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 1994 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 1994, as received and processed by CMS through July 1995. (See the variables in the Administrative Identification Record from **H_LATDCH** to the end). Individual bill records are supplied as part of this public use release for researchers who wish to look at Medicare bills in detail (i.e., the HOSPITAL BILL, the SNF BILL, the HOSPICE BILL, the HOME HEALTH BILL, the OUTPATIENT BILL, the PHYSICIAN/SUPPLIER BILL and the DURABLE MEDICAL EQUIPMENT BILL).

The utilization summary represents services rendered and reimbursed under fee-for-service in calendar year 1994. 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 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.

The variables pertaining to deductibles (Part A deductible, **H_INPDED**, Part B deductible, **H_PTBDDED**, and blood pints deductible, **H_BLDDED**) are always blank. This information is not consistently available from CMS's present files. An approximation can be derived from the individual bill records.

The variables pertaining to special coverage (lifetime reserve days, **H_RESDAY**, and psychiatric days, **H_PSYDAY**) are always blank. These benefits are applied to the beneficiary once in a lifetime, and they are decremented as they are used. At the current

time, CMS files contain a "current balance" of these benefit days rather than a history of their utilization.

Utilization summary

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 which 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, cancelling out something of 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.

Utilization summary

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. Admissions (**H_INPSTY**) were totaled by sorting the bills in chronological order, and counting the first admission in each sequence. 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. Admissions (**H_SNFSTY**) were totaled by sorting the bills in chronological order, and counting the first admission in each sequence. 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_SNFCCHG** 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 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 reimbursement (**H_HHARMB**) was summed from the variable **PROVPAY**.

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

Utilization is summarized by 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 number of claims, number of line items, submitted and allowed charges, reimbursement, office visits and office visit charges. All claims and individual line items (there can be up to 13 per claim) were counted and summed as (**H_PMTCLM**) and (**H_PMTLIN**). Submitted charges and allowed charges (**H_PMTTCH**) and (**H_PMTCHG**) were summed from **SUBCRG** and **ALLOWCRG** in the bill. Total reimbursement for Part B claims (**H_PMTRMB**) was summed from the variable **PAYAMT** in the bill.

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 which 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 for an individual 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 the 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 10 for people from the original and first supplemental samples. To maximize the usefulness of this public use release as a cross-sectional file, we have filled in this missing information from the original (Round 1, Round 4, or Round 7) interview. 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 Medicare record date of birth (**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 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_RACE**) 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 (**D_ETHNIC**) 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)

The respondent was allowed to define marital status categories (**SPMARSTA**); there was no requirement for a legal arrangement (for example, separated). (Initial interview variable)

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

SPHIGRAD: Education does not include education or training received in vocational, trade or business schools outside of the regular school system. This variable only includes years the sample person actually finished. If the sample person had earned a GED, the response was coded "high school--4th year". If the sample person said he or she earned a college degree in fewer than 4 years, the response was coded "college and graduate school--4 years". If the sample person attended school in a foreign country, in an ungraded school, under a tutor or under special circumstances, the nearest equivalent or the number of years of attendance was coded. (Initial interview variable)

INCOME: Income represents the best source or estimate of income during 1993. Round 9 represents the most detailed information for 1993 and is used when available. For individuals not completing Round 9 (that is, continuing sample people unavailable for that round and Round 10 supplemental sample people), the most recent information available was used. It should be noted that INCOME includes 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 below \$25,000 (or, conversely, \$25,000 or more).

Survey Health Status and Functioning Record (RIC 2)

The answers in the health status and functioning section of the questionnaire are a reflection of 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.

In the height measurement **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)

The sample person was asked to recall or estimate, not to measure or weigh himself or herself.

HYSTEREC: "Hysterectomy" includes partial hysterectomies. (Initial interview variable)

Use of other forms of tobacco, such as chewing tobacco, are 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. "Now" meant within the current month or so and not necessarily 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)

The answers about difficulty with various tasks (**DIFSTOOP**, **DIFLIFT**, **DIFREACH**, **DIFWRITE**, **DIFWALK**) reflect whether or not the sample person usually had trouble with these tasks, even if a short-term injury made them temporarily difficult.

The questions about various conditions (**OCARTERY**, **OCHBP**, **OCMYOCAR**, **OCCHD**, **OCOTHART**, **OCSTROKE**, **OCCSKIN**, **OCCANCER** **OCCLUNG**, **OCCOLON**, **OCCBREST**, **OCCUTER**, **OCCOROST**, **OCCCERVX**, **OCCBLAD**, **OCCOVARY**, **OCCSTOM**, **OCCKIDNY**, **OCCBRAIN**, **OCCTHROA**, **OCCBACK**, **OCCHEAD**, **OCCFONEC**, **OCCOTHER**, **OCDIABTS**, **OCARTHRRH**, **OCARTH**, **OCAARM**, **OCAFEET**, **OCABACK**, **OCANECK**, **OCAALOVR**, **OCAOTHER**, **OCMENTAL**, **OCALZHMR**, **OCPSYCH**, **OCOSTEOP**, **OCBRKHIP**, **OCPARKIN**, **OCEMPHYS**, **OCPPARAL** and **OCAMPUTE**) were coded 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)

IADLs and ADLs

"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.

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 Access to Care Record (RIC 3)Definitions applied to medical providers

Doctor - Medical doctors (M.D.) and doctors of osteopathy (D.O.). Chiropractors, nurses, technicians, optometrists, podiatrists, physician's assistants, physical therapists, psychologists, mental health counselors and social workers are not included. 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 - 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 - 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 - An organization that provides a full range of health care coverage in exchange for a fixed fee.

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.

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 samples

It should be noted in using data in this section that the questionnaire sequence on access to care for supplemental sample persons differs from that for continuing sample persons and may lead to apparent differences in expected number of responses to questions in the access to care codebook section. For example, continuing sample 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 sample 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.

USMCCHEK: The distinction in question US12 is between the doctor or doctor's office and the sample person or family. For example, if the check usually goes to the daughter, the answer would be coded "to the sample person."

USFINDMC: "Ever tried to find a doctor ..." refers to some type of active search. It does not refer to simply thinking or talking about it.

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.

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.
- 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 which 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 which are limited to certain illnesses or diseases such as cancer, stroke or heart attacks.
- Policies which 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_PHREL1 - D_PHREL5: The "Policy Holder or "Main insured person" is the member of the group or 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.

D_ANAMT1 - D_ANAMT5: 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, even though the sample person may not have held the policy for the full 12 months.

Survey Enumeration Record (RIC 5)

A household is defined as the group of individuals either related or unrelated who live together and share one kitchen facility. This may be one person living alone, a head of household and relatives only, or may include head of household, relatives, boarders and any other non-related individual living in the same dwelling unit.

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_HHUNREL**) 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 Identification Record (RIC 7)

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

Survey Interview Description Record (RIC 8)

This record was added for 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 10. 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_PROXRL**, **RRECHLP** and **D_IHLPRL** 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**. (NOTE: Interviewer remarks questionnaires are missing for of the community interviews).

"Sample person language problem" was given as a reason for the use of a proxy in 77 cases. More often, language problems were addressed without the use of a proxy. Interpreters were used in some cases, and Spanish-language versions of the questionnaires were used by bilingual interviewers 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 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 5 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.

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 16,119 people (longitudinal and Round 10 supplement sample persons) and can be used for making estimates of the population enrolled for Medicare for the whole of 1994.

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 (RIC X4, X3, X2)

The first set of non-zero longitudinal weights (RIC X4) applies to 8,237 individuals from the original sample who appeared in the CY 1991 (Round 1), CY 1992 (Round 4), CY 1993 (Round 7), and CY 1994 (Round 10) Access to Care releases (the weights for the remaining 7,882 individuals are zero). This set of weights can be used to subset the population in the study from the beginning of the survey and their experience for making comparisons between the 1991 and 1994 (as well as the 1992 and 1993) populations. This can be accomplished by match-merging the RIC K and the RIC X4 by BASEID and keeping all records with a non-zero weight in the RIC X4.

A second set of non-zero longitudinal weights (RIC X3) applies to 9,773 individuals (both original sample and Round 4 supplemental persons) who appeared in both the CY 1992 (Round 4), in the CY 1993 (Round 7), and in the CY 1994 (Round 10) Access to Care releases (the weights for the remaining 6,346 individuals are zero). This set of weights can be used to subset the continuing sample and their experience for making comparisons between the 1992 and 1994 (and 1993) populations. This can be accomplished by match-merging the RIC K and the RIC Y by BASEID and keeping all records with a non-zero weight in the RIC X3.

The third set of non-zero longitudinal weights (RIC X2) applies to 11,371 individuals (original sample and Round 4 supplemental and Round 7 supplemental persons) who appeared in both the CY 1993 (Round 7) and in the CY 1994 (Round 10) Access to care releases (the weights for the remaining 4,748 individuals are zero). This set of weights can be used to subset the continuing sample and their experience for making comparisons between the 1993 and 1994 populations. This can be accomplished by match-merging the RIC K and the RIC Y by BASEID and keeping all records with a non-zero weight in the RIC X2.

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 6 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 1994 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, 1994 and on or before December 31, 1994.
- Skilled nursing facility bills were included if the **admission or "from" date** fell on or after January 1, 1994 and on or before December 31, 1994.
- Home health agency and outpatient facility bills were included if the **"through" date** fell on or after January 1, 1994 and on or before December 31, 1994.
- Hospice bills were included if the **admission or "from" date** fell on or after January 1, 1994 and on or before December 31, 1994.
- Physician or supplier claims were included if the **latest "service thru" date** fell on or after January 1, 1994 and on or before December 31, 1994.
- Durable medical equipment (DME) claims were included if the **latest "service thru" date** fell on or after January 1, 1994 and on or before December 31, 1994.

A total of 2,934 (about 18.2 percent) of the sample people did not use Medicare reimbursed services in a fee-for-service setting in 1994; 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 13,185 individuals in the sample, we have captured bills meeting the date criteria, processed and made available by CMS through July 1995.

Medicare Current Beneficiary Survey
CY 1994 Access to Care

Edits

The use of Computer Assisted Person Interviewing (CAPI) significantly affects the data editing process. Many of the edits are performed as the responses are collected. Often, problems arising from miscommunications or data entry errors can be detected and corrected immediately. Also, since the computer software structures the interview, it prevents most "skip pattern" errors.

As survey information is collected, it is put into a database management system built into the CAPI software. During the interview and subsequent in-house review, the data in the database are subjected to two types of edits. First, logical relationship edits are performed between various segments of the database to ensure the integrity of the whole. Second, subject matter edits are performed to ensure the internal consistency of the data.

Logical relationship edits ensure that the database is sound by checking the links between segments. For example, every medical provider record in the provider segment must be linked to at least one respondent. The provider record is useless if the linkage does not exist.

Subject matter edits ensure the internal consistency of the data. These edits are of two types: those that result in changes to the database, and those that do not. Edits that result in changes to the database are not described in this document. The second group of consistency edits are the "no fix" edits. These edits serve as a warning that certain data are not consistent and cannot be made consistent with only the data and interviewers' notes for guides. These edits are described in Table 4.1. A list of the interviews that failed each edit follows the edit description.

Table 4.1: "No Fix" Internal Consistency Edits

NF10001 The number of children given in response to question IN14: "Including natural, adopted and stepchildren, how many living children do you have?" (Community component, Introduction); is less than the total number of people listed by name in the roster and identified as "son" or "daughter". (21 cases)

00208521	00210522	00210629	00217503
00220393	00221035	00221198	00225699
00230863	00231630	00234455	00237424
00238542	00241514	00243120	00249516
00253286	00263665	00266195	00266935
00269158			

NF07002 The sample person has indicated that he or she has trouble doing light housework, but has indicated no problem with heavy housework. (Community component, Instrumental Activities of Daily Living) (69 cases)

00011369	00019731	00020200	00023332
00028058	00036467	00046313	00052382
00075335	00080612	00082983	00090950
00091852	00094661	00096136	00098338
00099666	00100551	00102656	00105998
00107576	00122743	00150277	00150299
00153218	00156942	00158417	00159143
00163142	00170079	00182478	00185384
00186305	00187081	00189120	00193298
00195008	00196922	00206965	00207196
00208543	00209774	00211957	00214437
00215157	00224634	00225423	00226256
00228395	00230205	00230556	00235090
00235523	00237208	00241075	00241122
00241451	00245203	00249077	00250201
00254360	00255115	00259036	00260163
00262973	00266446	00268808	00269170
00269493			

Medicare Current Beneficiary Survey
CY 1994 Access to Care

Questionnaires

This section contains copies of the community and facility questionnaires that were administered during Round 10 of the Medicare Current Beneficiary Survey. Round 10 is the third 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.

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. Variable names are also indexed in the codebook.

Community Component

The community component is conducted in the home of the respondent. Since the community component of the survey was conducted using CAPI, the questionnaire actually exists 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.

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 1 for the original sample, Round 4 for the 1992 supplement, Round 7 for the 1993 supplement, Round 10 for the 1994 supplement, etc. 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 10, only persons in the longitudinal sample (that is, the original and first and second supplements) 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 10 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

This questionnaire is completed by the interviewer 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 the questionnaire was answered by proxy, 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.

Facility Questionnaire

The facility questionnaire is conducted conventionally using pen and paper 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 1994 Access to Care release is intended to serve as a "snapshot" of the sample person at one point in time during Round 10. 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 10.

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 4, 7, and 10) 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 4, 7, 10) 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.

Section 5: Questionnaire

Table 5.1 - Components of the Community Questionnaire

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

UPD	NAME/ADDRESS UPDATE
IN	INTRODUCTION
ENS*	ENUMERATION
--> EN	ENUMERATION
--> HI	HEALTH INSURANCE
UTS*	UTILIZATION SUMMARY
DU	DENTAL UTILIZATION AND EVENTS
ER	EMERGENCY ROOM UTILIZATION AND EVENTS
IP	INPATIENT HOSPITAL UTILIZATION AND EVENTS
IU	INSTITUTIONAL UTILIZATION
OP	OUTPATIENT HOSPITAL UTILIZATION AND EVENTS
HHS*	HOME HEALTH UTILIZATION SUMMARY
HH	HOME HEALTH UTILIZATION AND EVENTS
MP	MEDICAL PROVIDER UTILIZATION AND EVENTS
OM	OTHER MEDICAL EXPENSES UTILIZATION
PMS*	PRESCRIBED MEDICINE SUMMARY
PM	PRESCRIBED MEDICINE UTILIZATION
ST	CHARGE QUESTIONS (STATEMENT SERIES)
NS	CHARGE QUESTIONS (NO STATEMENT SERIES)
CPS*	CHARGE/PAYMENT SUMMARY
--> AC	PROVIDER PROBES/ACCESS TO CARE
--> SC	SATISFACTION WITH CARE
--> HS	HEALTH STATUS AND FUNCTIONING
--> US	USUAL SOURCE OF CARE
--> DI	DEMOGRAPHICS/INCOME (FOR SUPPLEMENTAL SAMPLE PEOPLE ONLY)
CL	CLOSING MATERIALS
--> IR	INTERVIEWER REMARKS

* 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.

Section 5: Questionnaire

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 1994 Access to Care

Sample Design and Guidelines for Preparing Statistics

This section opens with a description of the population covered by the 1994 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 non-response rates. Guidelines for preparing population estimates using full sample weights and variance estimates using replicate weights are then reviewed.

Medicare population covered by the 1994 public use data

The calendar year 1994 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 1994. This “always enrolled” population includes individuals enrolled on January 1, 1994 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, 1994; and
- 3) persons who dis-enrolled or died prior to the end of December 1994.

NOTE: A small number of sample people (153) included in this file died during 1994 subsequent to their Round 10 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).’

The “always enrolled” population concept was used for the CY 1991, CY 1992 and CY 1993 MCBS Access to Care releases for operational considerations, and is carried forth into this release for the same reasons. While it 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.

Section 6: Sample Design and Estimation

Table 6.1 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, gender, race using these three views: persons "ever-enrolled," persons enrolled as of the "mid-point of the year" (July 1), and persons "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 1992 file for that series.)

Section 6: Sample Design and Estimation

Table 6.1 1995 Medicare population, by gender, race and age

Gender Race	Age	Always July 1 Enrolled	Ever Midpoint	Ever Enrolled	
Total		34,598,440	36,621,680	38,484,500	
Females	Black	0-44	97,360	106,740	114,640
		45-64	187,140	204,440	218,120
		65-69	411,860	453,680	486,580
		70-74	392,720	399,080	405,880
		75-79	282,980	290,620	298,180
		80-84	202,420	209,320	217,220
		85+	198,140	209,860	223,500
	Non-black	0-44	406,900	446,860	479,220
		45-64	798,560	868,680	934,620
		65-69	4,320,480	4,801,080	5,199,200
		70-74	4,423,320	4,479,320	4,534,440
		75-79	3,453,600	3,519,260	3,589,100
		80-84	2,484,280	2,562,960	2,646,680
		85+	2,281,760	2,438,280	2,619,840
Males	Black	0-44	170,120	186,420	199,240
		45-64	229,740	248,500	263,560
		65-69	304,800	339,420	365,820
		70-74	264,060	272,280	280,660
		75-79	164,140	170,900	178,540
		80-84	96,060	101,140	107,440
		85+	71,000	76,700	83,280
	Non-black	0-44	681,280	742,960	793,300
		45-64	1,218,640	1,325,460	1,420,400
		65-69	3,626,600	4,049,360	4,399,440
		70-74	3,356,420	3,426,920	3,498,940
		75-79	2,281,320	2,353,500	2,430,320
		80-84	1,346,400	1,415,660	1,488,060
		85+	846,340	922,280	1,008,280
Female total		19,941,520	20,990,180	21,967,220	
Male total		14,656,920	15,631,500	16,517,280	
Black total		3,072,540	3,269,100	3,442,660	
Non-black total		31,525,900	33,352,580	35,041,840	

Based on March 1995 HISKEW files, inflated to 100 percent. "Always Enrolled" data are estimated.

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 and 10 the targeted populations included those individuals enrolled as of January 1, 1992,¹ as of January 1, 1993,² and as of January 1, 1994,³ respectively.

The targeted universe is divided into seven sampling strata based on age as of the midpoint (that is, July 1) of the year. 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 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 6.2.

Table 6.2 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

(Note: Due to the switch to a rotating panel design described earlier in the Introduction, it was necessary to interview roughly 16,000 sample persons in the fall 1994 round (Round 10) in order to meet the targets for complete annual utilization and expenditure data for both 1994 and 1995. This overlapping of panels will continue indefinitely under the rotating panel design for each fall interview session. See Table 6.3 for actual number of Round 10 completes by age stratum.)

Section 6: Sample Design and Estimation

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 mostly on considerations of convenience. The 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 group.

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 sample. In addition, for some PSU areas, the number of beneficiaries within the standard 5-percent sample is relatively small and the list of potential sample people can become exhausted. As a result, beginning in Round 10, samples are being drawn from 5-percent samples other than the standard 5-percent sample.

The MCBS sample is designed to be nearly self-weighting within the age strata. A systematic sampling scheme with random starts is employed. The use of random starts provides justification for the variance calculations described with the WESVAR procedures described later in this section.

Sample selection

A sample of 15,411 beneficiaries was selected in 1991 for Round 1 of the MCBS. This initial sample was representative of beneficiaries who were entitled on January 1, 1991. Round 1 interviews started in September of 1991, and the sample beneficiaries have been re-interviewed roughly every four months since then.

A supplemental sample of 2,410 beneficiaries was added to the sample for Round 4. The 1992 supplemental sample was primarily designed to include newly enrolled beneficiaries during the calendar year (from February 1991 through January 1992-see Endnote 1), as well as previously enrolled beneficiaries who were included for a coverage improvement component or to maintain the desired sample size in spite of the cumulative effects of deaths, emigration, and response rate losses. Several hundred persons pre-enrolled for some time in 1992 were inadvertently included.

The 1992 Access to Care questions were administered September through December 1992 as part of the Round 4 interview for the continuing sample, and as part of the initial interview of the supplemental sample.

An additional supplemental sample of 2,449 beneficiaries was added to the sample for Round 7. The 1993 supplemental sample included newly enrolled beneficiaries (from February through December

Section 6: Sample Design and Estimation

1992), as well as previously enrolled beneficiaries (that is, on or before January 1, 1992), to maintain the desired sample size.

The 1993 Access to Care questions were administered September through December 1993 as part of the Round 7 interview for the continuing sample, and as part of the initial interview of the 1993 supplemental sample.

An additional supplemental sample of 6,390 beneficiaries was added to the sample for Round 10. The 1994 supplemental sample included newly enrolled beneficiaries (from January through December 1993) as well as previously enrolled beneficiaries (again, through December 1992), to maintain the desired sample size.

The 1994 Access to Care questions were administered September through December 1994 as part of the Round 10 interview for the continuing sample, and as part of the initial interview of the 1994 supplemental sample.

Primary Sampling Units

The MCBS sample is spread across 107 primary sampling units (PSUs), which are metropolitan areas and clusters of non-metropolitan counties. Within the PSUs, the initial sample was concentrated in 1,163 clusters of ZIP code areas (5 digits). With the introduction of the 1992 and 1993 supplements, the number of sample ZIP code clusters expanded to 1,344 and 1,399, respectively. The area covered by the 1994 supplement included 1,432 clusters.

All four samples were selected from CMS's 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.

Response Rates

By Round 10, 68 percent of the original 1991 sample were still in a formal responding status (that is, either the SP was alive and still participating or had died after Round 1 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 1992 and 1993 supplemental samples, the corresponding figures were 71 and 73 percent, respectively. The 1994 supplemental sample (Round 10) had an initial response rate of 83 percent.

Section 6: Sample Design and Estimation

There were 8,237 interviews successfully completed at Round 10 with still-living members of the 1991 sample. For brevity, we refer to these 8,237 interviews as “live completes.” For the supplementary 1992, and 1993 samples there were 1,536, and 1,598 live Round 10 completes, respectively.

Completed interviews by sampling strata

Table 6.3 lists the number of completed interviews for Round 10 for the continuing (Rounds 1, 4 and 7) and supplemental (Round 10) samples by age strata.

Table 6.3 MCBS Round 10 completed interviews for the continuing and supplemental samples

Age group	- Round 10 Completed Interviews -		
	Continuing	Supplemental	Total
Total	11,371	4,748	16,119
0 - 44	854	397	1,251
45 - 64	1,044	414	1,458
65 - 69	1,890	847	2,737
70 - 74	1,905	781	2,686
75 - 79	1,849	799	2,648
80 - 84	1,950	746	2,696
85 +	1,879	764	2,643

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 low in the CY 1994 Access to Care release, but the analyst still needs to be aware of the missing data and be cautious about patterns of non-response.

Some of the missing data is attributable to the fact that some of the community interviews and all of the facility interviews are conducted by proxy. In other words, the respondent had no knowledge of the information sought on the sample person. In other situations the respondent 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.

There are other 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 10 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 **R10CWGT** (RIC X). R10CWGT 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 10. This weight is intended for use in cross-sectional statistics involving Round 10 by itself. Each weight is greater than zero for all 16,119 beneficiaries on the file. R10CWGT should be used to make estimates of the levels of access to care for the Medicare population alive in the fall of 1994.

The second set of weights is labeled **R110WGT** (RIC X4). It is intended for use in longitudinal statistics involving continuing people from the original 1991 and the 1992 supplemental samples who had Round 1, Round 4, Round 7, and Round 10 interviews. The longitudinal weight, **R110WGT**, does not apply to the Round 4, Round 7, and Round 10 supplemental persons. This weight should only be used when the CY 1991 (Round 1), CY 1992 (Round 4), CY 1993 (Round 7), and CY 1994 (Round 10) Access to Care releases have been merged together. Records must be merged at the beneficiary level, and only those beneficiaries who completed all four rounds (that is, have positive numeric R110WGT) should be included in the final file.

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 over the four-year period. R110WGT is greater than zero only for continuing persons in the original sample and the original sample; it is inapplicable (“.”) for the Round 4, Round 7, and the Round 10 supplemental samples. There are 8,237 beneficiaries who completed Round 1, Round 4, Round 7 and Round 10 interviews; these beneficiaries are identified as **D_SMPTYP**='91' (RIC K, Key Record).

The third set of weights is labeled **R410WGT** (RIC X3). These weights are intended for use in longitudinal statistics involving continuing sample people (from the original 1991 and the 1992 supplemental

Section 6: Sample Design and Estimation

samples) who completed interviews in Round 4, Round 7 and Round 10. The longitudinal weight, **R410WGT**, does not apply to the Round 7 supplemental sample or the Round 10 supplemental sample cases. This weight should only be used when the CY 1992 (Round 4), CY 1993 (Round 7), and CY 1994 (Round 10) Access to Care releases have been merged together. Records must be merged at the beneficiary level, and only those beneficiaries who completed all three interviews should be included in the final file.

This weight can then be used to make estimates of year-to-year changes from 1992 to 1994. **R410WGT** is greater than zero only for persons who completed Round 4, Round 7, and Round 10; it is inapplicable (“.”) for the Round 7 and Round 10 supplemental samples. There are 9,773 beneficiaries who completed Round 4, Round 7, and Round 10 interviews; these beneficiaries are identified as **D_SMPTYP=‘91’** or **D_SMPTYP=‘92’** (RIC K, Key record).

The fourth set of weights is labeled **R710WGT** (RIC X2). These weights are intended for use in longitudinal statistics involving continuing sample people (from the original 1991, the 1992 and the 1993 supplemental samples) who completed interviews in both Round 7 and Round 10. The longitudinal weight, **R710WGT**, does not apply to the Round 10 supplemental sample cases. This weight should only be used when the CY 1993 (Round 7) and CY 1994 (Round 10) Access to Care releases have been merged together. Records must be merged at the beneficiary level, and only those beneficiaries who completed both interviews should be included in the final file.

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. **R710WGT** is greater than zero only for persons who completed Rounds 7 and 10; it is inapplicable (“.”) for the Round 10 supplemental sample. There are 11,371 beneficiaries who completed both Round 7 and Round 10 interviews; these beneficiaries are identified as **D_SMPTYP=‘91’**, **D_SMPTYP=‘92’** or **D_SMPTYP=‘93’** (RIC K, Key record).

Although it is possible to create some cross-sectional estimates using **R110WGT**, **R410WGT**, or **R710WGT** and create longitudinal estimates using **R10CWGT**, 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 **R110WGT**, **R410WGT** or **R710WGT** are used for cross-sectional estimation, recently enrolled beneficiaries will not be represented. When **R10CWGT** is used for longitudinal estimation, positive weights will be given to cases without baseline data and the weights for long-term enrollees will be too small.

Variance Estimation (Using the Replicate Weights)

In many statistical packages, including SAS, the procedures for calculating variances assume that the data were collected in a simple random sample. Procedures of this type are not appropriate for calculating the variance for statistics based upon a stratified, unequal-probability, multistage sample such as the MCBS.

The replicate weights associated with the MCBS data can be used to create estimated standard errors for MCBS variables. Just as there are four full sample weights for Round 10, 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 **R10CS001** through **R10CS100** and may be found in the Cross-sectional Weights record (RIC X). The first set of replicate longitudinal weights are labeled **R110X001** through **R110X100** and may be found on the first Longitudinal Weights Record (RIC X4). The second set of replicate longitudinal weights are labeled **R410X001** through **R410X100** and may be found on the second Longitudinal Weights Record (RIC X3). The third set of replicate longitudinal weights are labeled **R710X001** through **R710X100** and may be found on the third Longitudinal Weights Record (RIC X2).

These replicate weights should be used for variance estimation. The user has three options for using the replicate weights. The first option is to use a software package called WesVar PC® that is available from Westat at no charge. The program can be downloaded from Westat's home page on the World Wide Web at: **WWW.WESTAT.COM**. Additional documentation is available from Westat and can be obtained by submitting a request to: **WESVAR@WESTAT.COM**.

Identification of weight variable and variables for analysis can be done using the WesVar PC® menus. To run WesVar PC® with MCBS data and weights, the method should be specified as Fay's method with a factor of 0.3.

The second 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 exact same statistic of interest formed using the corresponding 100 replicate weights. The estimated variance of X_0 is then simply:

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

Section 6: Sample Design and Estimation

The third option is for users who prefer to use alternate software such as SUDAAN® (Professional Software for SURvey DATA ANalysis for Multi-stage Sample Designs) to compute population estimates and the associated variance estimates. Two variables, SUDSTRAT and SUDUNIT, have been included in the cross sectional and longitudinal weight records to allow use of SUDAAN.

An example of using SUDAAN statements to compute an estimate of the proportion of the Medicare population (excluding newly enrolled after January 1995) that had an inpatient stay in 1995 is shown below:

```
PROC CROSSTAB DATA=dsn FILETYPE=SAS DESIGN=WR;
  NEST SUDSTRAT SUDUNIT / MISSUNIT;
  WEIGHT R13CWGT;
  SUBGROUP H_INPSW;
  LEVELS 2;
  PRINT NSUM ROWPER SEROW;
```

The data set dsn is the name of the file that contains the weights and statistics of interest. The variable H_INPSW is an inpatient switch indicator from the RIC A. Note that variables specified in the NEST, WEIGHT, and SUBGROUP statements have to be numeric and that SUBGROUP must run between 1 and the number of levels with no gaps.

Consistency with Medicare Program Statistics and with CY 1991, CY 1992, and CY 1993 Access to Care releases

In general, MCBS estimates are not consistent with Medicare program statistics such as tabulations of the HISKEW. There are several reasons for the inconsistencies. 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 expected to diminish somewhat due to recent efforts to modify CMS edit procedures that rejected records from SSA which had valid dates of death but other erroneous information and due to a recent replenishing of the EDB with SSA records to get nine-digit zip codes. Unrecorded deaths are still present on the EDB.

For cross-sectional estimates of the CY 1992 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 1993 or CY 1994.

Section 6: Sample Design and Estimation

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 re-weighting to \$5,035 per beneficiary after re-weighting. This downward bias was due to the presence of a relatively small number of individuals on the enrollment 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 re-weighting.

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 re-weighting 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 mis-sampled 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). Weighting procedures for the 1995 Access to Care release included a special adjustment to correct for the mis-sampling. 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.

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.