OVERVIEW OF HEALTH CARE FINANCING ADMINISTRATION DATA

TABLE OF CONTENTS

OVERVIEW OF HEALTH CARE FINANCING ADMINISTRATION DATA

The daily operation of the Medicare and Medicaid programs involves the processing, adjudication, and payment of individual claims for health care services. Disbursement of program funds is subject to numerous guidelines, schedules, and rules. As a result, extensive records are maintained on program participants, services, and payments. Administrative record-keeping requires the daily update of very large databases. While claims processing is distributed between many Medicare contractors and state Medicaid agencies, HCFA maintains a central repository of administrative records for program oversight, research, and evaluation purposes.

By linking, tabulating, sampling, and summarizing the records in the administrative databases, the Office of Information Services (OIS) creates analytic data files and program statistics required by various HCFA organizations for program management and policy development. The level of detail provided by the administrative records enables the creation of data files that support many analyses including patterns of service utilization and associated costs, and their variations across geographic areas, demographic groups, and diagnoses. The availability of person- and procedure-specific data also makes these files useful for epidemiological research applications.

The following text describes how Medicare and Medicaid data are organized and presented in the Data section of the guide. Included are descriptions of the categories used to group and describe the data and a table that provides an overview of files discussed in the guide.

DATA ORGANIZATION

The Data section of the Reference Guide contains database and file information. The databases and files are divided into data chapters and file groups. Organizing and grouping the data in this manner throughout the Reference Guide enables the reader to easily locate specific file information and to relate it to the HCFA data processing environment.

Data Chapters

The files discussed in the Reference Guide are grouped into ten data chapters. These groups are based primarily on data source and unit of analysis. For example, most of the files in the Claims and Utilization Data chapter have the claim as the unit of analysis and/or are derived from Common Working File System (CWF) claims data. The data chapters follow:

- **!** **Claims and Utilization Data**
- **!** **Enrollment and Eligibility Data**

!	**Combined Claims and Enrollment Data**
!	**Provider of Services Data**
!	**Provider Cost Report Data**
!	**Special Programs Data**
!	**Medicaid Data**
!	**Reference/Resource Data**
!	**Public Use Files Data**
!	**Beneficiary Encrypted Files Data**

The **Claims and Utilization Data** chapter discusses numerous analytic files that were created by HCFA to support Medicare program and policy evaluation and development. These files contain detailed records of Medicare claims submitted for payment as well as summarized claims data. As such, they provide a unique source of information on the utilization of health care.

Accurate and up-to-date records of each Medicare enrollee are required by HCFA for Medicare program administration. Eligibility of the individual receiving services and the status of beneficiary entitlement must be established each time a claim is submitted. HCFA operates an extensive enrollment data system to meet the information requirements of this administrative function. The **Enrollment and Eligibility Data** chapter describes files derived from the enrollment data system.

The **Combined Claims and Enrollment Data** file contains linked enrollment and utilization data. The file in this chapter contains beneficiary characteristics data and annual utilization data for a sample of Medicare beneficiaries. The file contains a history of annual utilization, regardless of utilization activity within a particular year.

The **Provider of Services Data** chapter describes three files. Two of the files contain information on institutional provider characteristics. One of these files contains data that are collected during the Medicare and Medicaid institutional provider certification process. The other file contains data used to calculate payments for all Medicare Prospective Payment System (PPS) hospitals. The remaining file contains summary information for physicians.

The **Provider Cost Report Data** chapter describes files containing data from cost reports that are submitted to HCFA by Medicare-certified facilities. The files contain specific financial and statistical cost report data such as facility characteristics and costs and charges by cost center.

The **Special Programs Data** chapter consists of files containing information collected for a particular group of Medicare beneficiaries or providers. This chapter discusses the End Stage Renal Disease (ESRD) Program Management and Medical Information System (PMMIS) Database and the Medicare Current Beneficiary Survey (MCBS) Files.

The files discussed in the **Medicaid Data** chapter contain Medicaid eligibility, expenditure, and recipient data. Also included are state Medicaid program characteristic data, and Medicaid drug utilization and product description data.

The **Reference/Resource Data** chapter presents a collection of utility files used by HCFA in conjunction with administrative and analytic files. The files contain various indices, standards, and coding systems.

The **Public Use Files Data** chapter describes files that have been edited and stripped of all information that could be used to identify individuals. These files are the primary source of data for federal agencies outside of HCFA, government contractors, academic researchers, and commercial enterprises that are not permitted access to individual identification information.

The **Beneficiary Encrypted Files (BEFs) Data** chapter includes files that have been encrypted, ranged or blanked. BEFs require a signed Data Use Agreement (DUA) before the data can be released. These files are another source of data for customers (federal agencies outside of HCFA, government contractors, academic researchers, and commercial enterprises) who are not permitted access to individual identification information, but have projects determined to be in the government's interest and projects that require a level of detail not available in the PUFs.

File Groups

The files within five of the data chapters are further divided into file groups that organize the files by system or function. The file groups are presented below, under the corresponding data chapter. In other data chapters where no file groups exist, there were no logical groupings of the files, or there were too few files to establish subgroups.

! Claims and Utilization Data
- National Claims History Files
- Standard Analytic Files
- Stay Records Files
- Part B Medicare Files
- Other Utilization Files

! Enrollment and Eligibility Data
- Enrollment Database Files
- Other Enrollment Files

! Medicaid Data
- Medicaid Eligibles, Claims, and Utilization Files
- Medicaid Drug Rebate Files
- Medicaid and Children's Health Insurance Budget and Expenditure Files

!   Public Use Files Data
-      Cost Limits
-      Cost Reports
-      Enrollment
-      Payment Rates Institutional Providers
-      Payment Rates Non-Institutional Providers
-      Providers
-      Utilities/Miscellaneous
-      Utilization
-      Medicaid

!   Beneficiary Encrypted Files Data
-      Enrollment
-      Standard Analytical Files (SAFs)
-      Expanded Modified MEDPAR - Hospital (National) File
-      Expanded Modified MEDPAR - Hospital (State) File
-      Expanded Modified MEDPAR - Skilled Nursing Facility
-      Physician Sample File (Formerly BMAD Provider File)
-      5% Sample Durable Medical Equipment (DME)

File Discussions

The information presented for most files is organized in a consistent manner under the categories described below. These categories correspond to the icons introduced in Orientation for the Reader. In some file discussions, the use of the categories is dropped to minimize exact duplication of information. In such cases, references to other file discussions containing relevant information are included. The Reference/Resource Files, Public Use Files (PUFs) and Beneficiary Encrypted Files (BEFs) are described briefly, and again reference other data chapters where appropriate.

By using the tables of contents and icons, the reader will be able to find information regarding a specific file easily and quickly. To establish a context for the file discussions, the reader is encouraged to review the introductions to the data chapters and file groups.

File Creation

File Creation includes a discussion of the data collection method and file population or sample. The process of creating the file is discussed, including cut-off points for including data, date of file maturation, and Quality Assurance (QA) measures executed when the file is created. For files derived from other files, the parent file is referenced and the method of derivation is discussed.

Issues of file availability, including years for which the file is available and length of the file creation cycle, also are addressed.

File Maintenance
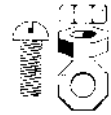
A description of the file update cycle and QA measures involved in updating and maintaining the file are provided in File Maintenance. The information in File Maintenance is closely related to that in File Creation. Text may be repeated across these categories or they may reference one another.

File Structure and Usage

File Structure and Usage describes the type of information contained in the file or database. A general description of the data elements that comprise the file and the unit of analysis are also included. Known data limitations at the file level are discussed when applicable, as are general uses of the file to both the research and policy making communities.

Data Structure and Usage
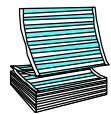
Data samples, data availability, and data usage are addressed under Data Structure and Usage. The type of variables on each record, total number of records in each file, and a discussion of known data limitations at an element level are provided. The information in Data Structure and Usage is closely related to that in File Structure and Usage. Text may be repeated across these categories or they may reference one another.

Methods of Access

Methods of Access discusses the various means available for file access. Many files can be accessed by one or more of HCFA's online access systems such as the Decision Support Access Facility (DSAF), Mainframe HCFA Customer Information System (HCIS) or Enrollment Database Workbench (EDBW). This part of the file discussion highlights these systems and addresses their current functions. Other methods of accessing the files, such as through flat files or SAS datasets, are also discussed.

When applicable, the Methods of Access section identifies the HCFA files available for purchase as a BEF or a PUF.

Migration of Data
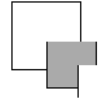
Continuity of data in the changing HCFA data processing environment is addressed in Migration of Data. For example, if the method of data storage has changed for a file, the ancestor file is identified, and the relationship between the former and new file is explained. This section also addresses future plans for data storage.

List of Data Elements

Data element lists for the files discussed within a chapter will appear in the future on the HCFA web site at www.hcfa.gov. These lists include the full names of the elements in the file, descriptions, and coding categories. The data element lists are generated from the HCFA Data Dictionary which is maintained by OIS, EDG, Division of Data Services (DDS).

OVERVIEW OF DATA USERS REFERENCE GUIDE FILES

The following summary table presents the key characteristics of each file discussed in the guide except the Reference/Resource Data files, PUFs and BEFs. Brief discussions of the Reference/Resource Data files, PUFs and BEFs can be found in their respective chapters. The table provides users with an overview of the files and their contents, and serves as a quick reference to a file that might meet specific data needs. The table is organized by data chapter, file group, and then file. For each file, the unit of analysis, data sample, years available, and file creation cycle are noted.

OVERVIEW TABLE OF FILES

CLAIMS AND UTILIZATION DATA				
FILE GROUP or *File Name*	**Unit of Analysis**	**Sample**	**Years**	**Frequency of File Creation**
NATIONAL CLAIMS HISTORY FILES				
National Claims History (NCH) 100% Nearline File	Claim as submitted by CWF site; adjustments and/or interim billing may result in multiple claims for a service/stay	All claims for 100% of Medicare beneficiaries	Nearline: 1991-present	Monthly and annually
STANDARD ANALYTIC FILES				
Inpatient 100% Standard Analytic File (SAF)	Final action claim; all adjustments resolved	All Inpatient claims for 100% of Medicare beneficiaries	1991-present	Annual files produced quarterly
Outpatient 100% Standard Analytic File (SAF)	Final action claim; all adjustments resolved	All Outpatient claims for 100% of Medicare beneficiaries	1991-present	Annual files produced quarterly
Home Health Agency (HHA) 100% Standard Analytic File (SAF)	Final action claim; all adjustments resolved	All HHA claims for 100% of Medicare beneficiaries	1991-present	Annual files produced quarterly
Hospice 100% Standard Analytic File (SAF)	Final action claim; all adjustments resolved	All Hospice claims for 100% of Medicare beneficiaries	1991-present	Annual files produced quarterly
Skilled Nursing Facility (SNF) 100% Standard Analytic File (SAF)	Final action claim; all adjustments resolved	All SNF claims for 100% of Medicare beneficiaries	1991-present	Annual files produced quarterly
Clinical Laboratory 100% Standard Analytic File (SAF)	Final action claim; all adjustments resolved	All clinical laboratory claims for 100% of Medicare beneficiaries	1991-present	Annual files produced quarterly
Durable Medical Equipment (DME) 100% Standard Analytic File (SAF)	Final action claim; all adjustments resolved	All DME claims for 100% of Medicare beneficiaries	1991-present	Annual files produced quarterly

CLAIMS AND UTILIZATION DATA				
FILE GROUP or *File Name*	**Unit of Analysis**	**Sample**	**Years**	**Frequency of File Creation**
*5% Sample Beneficiary Standard Analytic File (SAF)**	Final action claim; all adjustments resolved	All claims for 5% of Medicare beneficiaries	1991-present	Annual files produced quarterly
STAY RECORDS FILES				
Medicare Provider Analysis and Review (MEDPAR) File	Hospital or SNF Stay	Inpatient/SNF claims for only 20% of Medicare beneficiaries from 1979-1983 and for 100% of beneficiaries from 1984-present	1984-present	Quarterly and annually
PART B MEDICARE FILES				
*Physician/Supplier Procedure Summary File**	Unique combinations of carrier, procedure code, first and second modifiers, type of service, place of service, specialty, and locality	100% of physician/ supplier procedures that occur in Part B Medicare claims	1991-present	Quarterly
*Physician Sample File**	Claim line item	Random sample of physicians using all line items	1991-present	Annually
OTHER UTILIZATION FILES				
Provider Summary File	Provider	All claims for 100% of Medicare beneficiaries	1987-present	Monthly
Physician Summary File	Physician	Charges, payments, and services by provider category	1991- present	Annually

*These files are successors to the BMAD files. Further information about the BMAD successor files can be found in the Migration of Data section for each BMAD file discussion in the Claims and Utilization Data chapter.

ENROLLMENT AND ELIGIBILITY DATA				
FILE GROUP or *File Name*	**Unit of Analysis**	**Sample**	**Years**	**Frequency of File Creation**
ENROLLMENT DATABASE FILES				
Enrollment Database (EDB)	Enrollee	Everyone ever enrolled in Medicare	Beginning of Medicare program to present	Daily
5%+ Enrollment Database (EDB)	Enrollee	5% of population ever enrolled in Medicare and 100% of ESRD enrollees	Beginning of Medicare program to present	Monthly
OTHER ENROLLMENT FILES				
Third Party Master (TPEARTH) File	Beneficiary	100% of Medicare beneficiaries for whom Part A and/or Part B premiums are paid by a Third Party	1978-present	Monthly
Group Health Plan (GHP) Master File	Beneficiary	100% of Medicare beneficiaries who have ever been enrolled in an MCO since 1983	1983-present	Daily Flat File-Monthly
Health Insurance Skeleton Eligibility Write-off (HISKEW) File	Beneficiary	100% of Medicare beneficiaries who have ever been enrolled	1989-present	Quarterly
Denominator File	Beneficiary	100% of entitled Medicare beneficiaries for specified year	1984-present	Annually
Name and Address File	Beneficiary	100% of Medicare beneficiaries	Current data	Ad hoc

COMBINED CLAIMS AND ENROLLMENT DATA				
FILE GROUP or *File Name*	**Unit of Analysis**	**Sample**	**Years**	**Frequency of File Creation**
Continuous Medicare History Sample (CMHS) File	Beneficiary	All claims for 5% of Medicare beneficiaries	1974-present	Annually

PROVIDER OF SERVICES DATA				
FILE GROUP **or** *File Name*	**Unit of Analysis**	**Sample**	**Years**	**Frequency of File Creation**
Provider of Services (POS) File	Provider	All Medicare and Medicaid institutional providers	1987-1989; 1991-present	Annually; Quarterly
Provider Specific File	Provider	PPS-eligible hospitals	1983-present	Quarterly
Medicare Physician Identification Eligibility Registry (MPIER) File	Physician	Medicare physician identification information	1993-present	Quarterly

		PROVIDER COST REPORT DATA		
FILE GROUP or *File Name*	**Unit of Analysis**	**Sample**	**Years**	**Frequency of File Creation**
Hospital Minimum Datasets	Hospital	All Medicare-certified hospitals and hospital/health care facility complexes	1981-present	Quarterly
*Exempt Hospital and Excluded Unit Dataset*s	Hospital	All PPS-exempt hospitals and PPS hospitals with excluded units	1985-present	Annually
Hospital Part B Datasets	Hospital	Four types of Part B services for all Medicare-certified hospitals: Ambulatory Surgical Centers, Outpatient Radiology, Other Outpatient Diagnostics, and "All Other"	1989-present	Quarterly
Hospital Capital Datasets	Hospital	All Medicare-certified hospitals	1985-present	Quarterly
Hospital Worksheet A Datasets	Hospital	All Medicare-certified hospitals	1992-present	Quarterly
Hospital Cost Report Systems Master File	Hospital	All Medicare-certified hospitals	1992-present	Quarterly
Skilled Nursing Facility (SNF) Datasets	SNF	All Medicare-certified Skilled Nursing Facilities	1989-present	Quarterly
End Stage Renal Disease (ESRD) Datasets	Renal facility	All Medicare-certified free-standing and hospital-based renal facilities	1989-present	Quarterly
Home Health Agency (HHA) Datasets	HHA	All Medicare-certified Home Health Agencies	1994-present	Quarterly

SPECIAL PROGRAMS DATA

FILE GROUP or *File Name*	Unit of Analysis	Sample	Years	Frequency of File Creation
Renal Beneficiary and Utilization System (REBUS)/ Management and Medical Information System (PMMIS) Database for the End Stage Renal Desease (ESRD) Program	ESRD Beneficiary or Renal Provider	100% of Medicare beneficiaries with ESRD and 100% of Medicare-approved renal providers	1973-present	Beneficiary data: monthly and annually; Provider data: ad hoc
Medicare Current Beneficiary Survey (MCBS) Files	Beneficiary	Statistically representative sample of Medicare beneficiaries-- approximately 13,000 aged and 2,500 Medicare beneficiaries with disabilities	1991-present	Annually

MEDICAID DATA

FILE GROUP or *File Name*	Unit of Analysis	Sample	Years	Frequency of File Creation
MEDICAID ELIGIBLES, CLAIMS, AND UTILIZATION FILES				
Medicaid Valid Tapes File--Eligible File	Medicaid eligible individual	100% of Medicaid eligible individuals in MSIS participating States	1985-present	Quarterly
Medicaid Valid Tapes Files-- Claim Inpatient (Claim-IP) Claim Long-term (Claim-LT) Claim Other (Claim-OT) Claim-RX	Claim	All claims for 100% of Medicaid recipients in MSIS participating States	1985-present 1999-present	Quarterly
Medicaid Personal Summary Record File	Recipient	Summary of Medicaid eligibility claims for 100% of Medicaid eligibles in MSIS participating States	1985-1998	Quarterly

MEDICAID DATA

FILE GROUP or *File Name*	Unit of Analysis	Sample	Years	Frequency of File Creation
State Medicaid Research Files-- (SMRFs) Claim Inpatient (Claim-IP) Claim Long-term (Claim-LT) Claim Other (Claim-OT) Claim Drug (Claim-Rx)	Claim	All claims for 100% of Medicaid recipients in MSIS participating States	1992-1994	Annually
Person Summary Record File	Recipient	Summary of Medicaid eligibility and claims for MSIS participating States by calendar year. DB2 Structure will replace summary and SMRF in 1999.	1992-1994	Annually
Medicaid Tape Option HCFA Form 2082 Totals File	State	All Medicaid recipients, services, and payments in all States	1987-present	Quarterly
MEDICAID DRUG REBATE FILES				
Medicaid Drug Rebate Product Description File	11-Digit National Drug Code	All Medicaid covered drugs	1991-present	Quarterly
Medicaid Drug Rebate Utilization File	11-Digit National Drug Code and State	Medicaid drug utilization for all States	1991-present	Quarterly
Medicaid Budget and Expenditure System (MBES) File	State	Medicaid budget and expenditure data for participating states	1989-present	Quarterly
		Children's Budget and Expenditure Data	1998-present	Quarterly

OFFICE OF INFORMATION SERVICES (OIS) FACILITIES FOR ACCESSING ENROLLMENT AND UTILIZATION DATA

Several facilities for accessing a variety of enrollment and utilization files are available to HCFA Data Center (HDC) users. The OIS Enrollment and Utilization Access Facilities Table at the end of this chapter provides an overview of these access facilities. For each facility, the table indicates when the

facility should be used, search attributes, and files that can be accessed. Brief descriptions of the access facilities listed in the table are also provided.

Decision Support Access Facility (DSAF)

The first access facility listed in the OIS Access Facilities Table, DSAF, provides users with a single access path to a wide array of Medicare data. Through DSAF, users can extract Medicare enrollment, entitlement, and utilization databases and files. DSAF provides easy-to-use screens and online instructions. Data selections can be based on a finder list, specific data fields, or a percentage sample. DSAF is divided into logical groupings referred to as "legs." Each leg provides access to different groups of files. A Summary Figure of the files available in each leg under DSAF can be obtained by contacting the OIS, EDG, Division of Information Distribution (DID) at 410-786-0159. DSAF is available to authorized HDC users. A written request must be submitted to the OIS, EDG, Division of Data Liason and Distribution (DDLD).

Enrollment Database Workbench (EDBW)

The EDBW is a menu-driven retrieval system that allows the creation of customized user-views containing enrollment and entitlement information. The EDBW can be used for access to current, complete Medicare beneficiary-specific elements such as personal characteristics, entitlement dates, and types of coverage.

Beneficiary Enrollment Retrieval (BERT) System

The BERT System is an online query facility that provides access to Medicare beneficiary demographic and enrollment and entitlement data. BERT is used to view specific beneficiary Enrollment Database (EDB) records.

Part B Extract and Summary System (BESS)

BESS is a menu-driven query system that provides multiple-path access to non-beneficiary specific physician/supplier claims data that have been summarized at the procedure code level. Additionally, BESS provides online access to Part B carrier information (e.g., carrier numbers, addresses, localities) and all procedure and modifier codes (e.g., Current Procedural Terminology-4 [CPT-4] codes, HCFA alpha-numeric codes, carrier local codes).

HCFA Customer Information System (HCIS)

The HCIS is a multi-dimensional software application that provides an easy-to-use access path for non-programmers to manipulate Medicare data into information. This software application is commonly referred to as the General Purpose Front End (GPFE) to the system. One aspect of HCIS

provides Graphical User Interface (GUI) forms (views) and reports on the different types of Medicare services. Currently, the GPFE provides access to the following types of provider data: Inpatient, Outpatient, Skilled Nursing Facility (SNF), Home Health Agency (HHA), Hospice, Physician Services, Non-Physician Practitioner, Ambulance Services, Enrollment, Durable Medical Equipment (DME), and Clinical Laboratory Data. HCIS uses as source data the National Claims History (NCH) database and the Standard Analytical Files (SAFs). Data is segmented by calendar year based on date of service.

A second aspect of HCIS is the mainframe access path to views containing key claims level data on individual beneficiaries, providers and beneficiary/provider data combinations. Views currently provided include SNF, HHA, Hospice, Outpatient, Clinical Lab and DME. The views provide instant access to a ranking of beneficiaries by provider according to service utilization and claims reimbursement. The views supplement the functionality provided under the HCIS General Purpose Front End.

Menu-driven Access to the 100% Nearline Claims File (MANRLINE) System

MANRLINE is an online, menu-driven system for submitting retrieval requests for NCH 100% Nearline File data. MANRLINE allows users selections of groups of NCH claims records based on the values of one or more variables such as Health Insurance Claim (HIC) number, beneficiary state, beneficiary county, provider type, provider number, Unique Physician Identification Number (UPIN), diagnosis, Diagnosis Related Group (DRG), procedure code (ICD-9-CM and HCFA Common Procedure Coding System (HCPCS)), pricing locality, revenue center, and claim years. MANRLINE requests are processed bi-monthly. Each MANRLINE cycle encompasses four years of sequential data; i.e., 1991-1994, 1992-1995 or 1993-1996.

Office of Information Services (OIS)
Enrollment and Utilization Access Facilities Table

ACCESS PATHS	WHEN TO USE			SEARCH ATTRIBUTES					ACCESSIBLE FILES					
	COUNT INFO	EXTRACT	FINDER FILE CREATION	NUMERIC SEARCH HIC/SSN/RRB	ALPHA NAME SEARCH	GEOGRAPHIC SEARCH	DEMOGRAPHIC SEARCH	EXTENDED SELECTION CAPABILITY	EDB DATA	NCH DATA	SUM NCH DATA	SAF FILES	CARRIER INFO DB	HCPCS CODES DB
DSAF		%	%	%	%	%	%	%	%	%		%		
EDBW		%		%					%					
BERT				%					%					
BESS		%				%		%			%		%	%
HCIS	%	%	%	%				%		%		%		
MANRLINE		%	%	%				%		%				