TABLE OF CONTENTS

CLAIMS AND UTILIZATION DATA 1
INTRODUCTION 1
The Creation of Claims and Utilization Files 2
NATIONAL CLAIMS HISTORY FILES 4
National Claims History (NCH) 100% Nearline File 4
STANDARD ANALYTIC FILES (SAF) 9
Inpatient 100% SAF 11
Outpatient 100% SAF 12
Home Health Agency (HHA) 100% SAF 12
Skilled Nursing Facility (SNF) 100% SAF 12
Hospice 100% SAF 12
Clinical Laboratory 100% SAF 12
Durable Medical Equipment (DME) 100% SAF 13
5% Sample Beneficiary SAF 13
STAY RECORDS FILES 13
Medicare Provider Analysis and Review (MEDPAR) File 14
PART B MEDICARE FILES 17
Physician/Supplier Procedure Summary (PSPS) File 18
Physician Sample File 21
OTHER UTILIZATION FILES 22
Provider Summary File 23
Physician Summary File 24

CLAIMS AND UTILIZATION DATA

INTRODUCTION

Processing claims for Medicare health insurance benefits is fundamental to the operation of the Medicare program. HCFA is responsible for ensuring that payments are made for medically appropriate and covered services. HCFA also ensures that these services are rendered to eligible beneficiaries by qualified providers. Extensive records are generated for and by this administrative function. The detailed claims records submitted for Medicare covered services provide a unique source of information on health care utilization and costs. From these records, numerous analytic files are created by HCFA to support Medicare program and policy development and evaluation, health care analyses, and clinical and epidemiological research.

The data for claims and utilization files originate with a claim submitted by a physician, hospital, or other provider. In the last few years, the process by which claims are submitted and paid has undergone changes that have significantly improved data availability and quality.

The Claims and Utilization Data files contain extensive utilization information at various levels of summarization for a variety of providers and services. Depending on the level of summarization, the unit of analysis for a given file of claims records might be the claim, hospital stay, or procedure. Some files contain data from a single type of provider, such as a hospice, while others contain data from more than one type of Medicare provider.

The Claims and Utilization Data chapter consists of five file groups:

- National Claims History (NCH) Files
- Standard Analytic Files (SAFs)
- Stay Records Files
- Part B Medicare Files
- Other Utilization Files

The NCH Files contain claim records for both institutional (Inpatient/SNF/Outpatient, HHA and Hospice) and non-institutional (Physician/Supplier & DME) data. The SAFs contain claims records in final action form, with all adjustments resolved. The Stay Records File group summarizes all inpatient services rendered to a beneficiary from admission to a hospital through discharge, or from admission to a Skilled Nursing Facility (SNF) (beneficiary may still be a patient in the SNF). The Part B Medicare Files group discusses the historic Part B Medicare Annual Data (BMAD) Files and their successors which contain data on the utilization and prices of physician services and certain other medical services, equipment, and supplies covered by the Medicare Supplementary Medical

Insurance (SMI) plan. Finally, the Other Utilization Files group contains claims and utilization files that
pre-date the advent of the Common Working File (CWF) (as do the historic BMAD files and the Stay Records Files).

Each of these file groups and the files within them are discussed in detail following background information on the creation of claims and utilization files.

The Creation of Claims and Utilization Files

The CWF is a Medicare Part A and Part B benefit coordination and claims validation system. Under CWF, the country is divided into nine distinct processing sectors. Each sector has a designated contractor "host" site and a number of Fiscal Intermediary (FI) or carrier processing contractors. Each beneficiary is assigned to one and only one sector. The host site maintains an entitlement and utilization database that includes all Part A and Part B utilization data for each Medicare beneficiary in its sector. FIs and carriers do not send claims directly to HCFA. Instead, they interact with the host site. FIs and carriers process claims submitted by providers, submit claims to the host site for prepayment review and authorization, and then act on host site authorization to pay claims. Further information about CWF claims processing can be found in the Overview of the Health Care Financing Administration chapter in the Orientation Section.

Prior to the CWF, claims were referred to as bills and HCFA stored institutional utilization and physician/supplier services data in separate files. The institutional utilization data were stored in five bill files: inpatient, Home Health Agency (HHA), hospice, outpatient, and SNF. The bill files were used to create other files such as the Medicare Provider Analysis and Review (MEDPAR) File. To supplement its physician/supplier utilization data, HCFA required all carriers to submit BMAD Files. BMAD Files contained sample and summary physician/supplier claims data, including claim line item detail for a sample of beneficiaries. More information on the BMAD Files can be found in the Part B Medicare Files group discussion in this chapter.

Under the CWF, HCFA maintains claim level data for both institutional and non-institutional services. Carriers no longer submit payment records or BMAD Files. Physician/supplier claims data contain line item detail for every service listed on the claim. CWF claims records are the data source for most claims and utilization files.

The following table shows the type of claims data contained in each claims and utilization file. The "Institutional and Non-institutional" column lists files that include data from both institutional and physician/supplier claims. The "Inpatient/SNF," "Outpatient," "HHA," and "Hospice" columns list the files that contain data only from each of those types of institutional claims. The last column presents files created from non-institutional claims data only. More information on each file in the table is provided in the individual file discussions.

Files containing claims data that are discussed in other chapters of the guide such as the Medicare Current Beneficiary Survey (MCBS) File and the Continuous Medicare History Sample (CMHS) File, are not included in this table.

CLAIMS AND UTILIZATION FILES					
Institutional and Non-Institutional Claims Data	**Institutional Claims Data Only**				**Non-institutional Claims Data Only**
	Inpatient/ SNF	**Outpatient**	**HHA**	**Hospice**	
! NCH 100% Nearline File ! 5% Sample Beneficiary SAF	! MEDPAR ! Inpatient 100% SAF ! SNF 100% SAF ! Provider Summary File	! Outpatient 100% SAF	! HHA 100% SAF	! Hospice 100% SAF	! Physician/ Supplier Procedure Summary File ! Annual Physician Fee Schedule Transition Payment Amount File ! 5% Beneficiary Physician/ Supplier Data (subset of 5% Sample Beneficiary SAF) ! Clinical Laboratory 100% SAF ! DME 100% SAF ! Physician Sample File

NATIONAL CLAIMS HISTORY (NCH) FILES

The first file group in the Claims and Utilization Data chapter is the NCH 100% Nearline File. The NCH 100% Nearline File contains all institutional and non-institutional claims from the CWF. It provides records of every claim submitted, including all adjustment claims (which amend claims originally submitted with errors). For analyses that do not require raw claims data, the SAFs offer final action claims data in the form of a netted (adjustment-resolved) file.

National Claims History (NCH) 100% Nearline File

The NCH 100% Nearline File contains all CWF-processed institutional provider and non-institutional claims record data.

File Creation

The NCH 100% Nearline File is created from claims records submitted to HCFA by CWF host sites. A description of the steps leading up to the creation of the NCH File can be found in the Overview of the Health Care Financing Administration chapter in the Orientation Section. The NCH 100% Nearline File is the only file that provides access to 100 percent of physician/supplier claims data.

Daily transmissions of claims data from the nine CWF host sites are released in weekly batches for uploading into the NCH 100% Nearline File. Each month four to five weekly files are merged and 99 files based on state of beneficiary residence are created (larger states are divided into several segments based on Health Insurance Claim (HIC) sequence segmentation). After 18 months, all files are merged to create an annual file which is approximately 98% complete at that time. Claims continue to trickle into the Nearline File for the next 30 months until the file is "closed" for that particular year. For example, the file for calendar year 1993 was closed in December 1996.

Enterprise Databases Group (EDG) eliminates some fields from the claims data transmitted to HCFA by the CWF hosts. Most of these fields contain CWF administrative information. A few data fields are added to NCH 100% Nearline records. Some of the added fields contain administrative data (e.g., HCFA claim process date), and others are derived variables (e.g., NCH Beneficiary Discharge Date--on an inpatient and HHA claim, the date the beneficiary was discharged from the facility or died).

Claims data are subjected to the NCH/CWF Medicare Quality Assurance System (NCH/CWF MQAS). NCH/CWF/MQAS is a three tier sequential data editing process. The first two tiers are applied to daily transmittals of CWF data from the host sites. The third tier is applied to NCH 100% Nearline data.

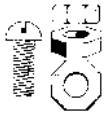
Tier 1: Daily files are examined to ensure that they have been properly and completely transmitted.

Tier 2: Edits performed at the CWF host sites are replicated to ensure adherence to contractor's specifications. A five percent sample of data is subjected to a more intensive screening process.

Tier 3: A sample of institutional and physician/supplier claims is extracted from the NCH 100% Nearline File and analyzed to identify Medicare contractors that exhibit unusual or inconsistent coding practices and to indicate data elements that are generally reliable or particularly unreliable. Information on unusual coding patterns is made available to HCFA components.

The NCH 100% Nearline File includes data from 1991 to the present stored on tape at the HCFA Data Center (HDC). The annual files are organized by claim through date. Current year data are available to the user within one month of being received by HCFA from the CWF host sites.

File Maintenance



The NCH 100% Nearline File is updated monthly with records received from the CWF host sites as described in File Creation.

File Structure and Usage

The NCH 100% Nearline File is divided into six record types. Four of the record types are for institutional claims and two are for non-institutional claims. The institutional record types include inpatient/SNF, outpatient, HHA, and hospice. The non-intitutional claims record types include physician/supplier and Durable Medical Equipment Regional Carrier (DMERC) processed claims. The DMERC-processed claims record was phased in beginning with October 1993 with complete implementation by July 1994. Prior to the new DMERC processed claims record type, the DME data was reported as a physician/supplier claim. Migration of Data provides more information about the DMERC-processed claims record and other changes to the NCH record format.

The NCH 100% Nearline File contains detailed information on beneficiary demographics, claims processing, diagnoses, procedures, and payment. It is the only file that contains line item detail for 100 percent of physician/supplier claims data. Prior to 1991 when the CWF was implemented, the BMAD Beneficiary File was the only source of line item detail from physician/supplier claims. The

BMAD Beneficiary File, however, contained data on only a five percent sample of all Medicare beneficiaries plus 100 percent of beneficiaries with End Stage Renal Disease (ESRD).

The unit of analysis in the NCH 100% Nearline File is the claim. More than one claim may be associated with a Medicare service such as a hospital stay or visit to the doctor. For example, a long hospital stay may generate one or more interim claims and a final claim. In addition, a claim containing incorrect information may be superseded by an adjustment claim, or a claim may be denied. The NCH 100% Nearline File contains all CWF-processed claims including adjustment, interim claims, and physician/supplier denials.

Data Structure and Usage

The NCH 100% Nearline File consists of six billion records: over 806 million records were created in 1998 alone. Annual NCH 100% Nearline files are organized by year of service based on claim through date. Annual files are segmented into 99 subgroups based on state of beneficiary residence and sequenced by the Beneficiary Claim Locator Number (Beneficiary Claim Account Number and NCH Category Equatable Beneficiary Identification Code (BIC)). The NCH 100% Nearline File data are stored on tape at the HDC in sequential flat files comprised of variable length records.

With the Version H upgrade, the NCH 100% Nearline File includes approximately 375 variables. Each record type contains a fixed portion consisting of variables that can occur only once per record and trailer groups consisting of variables that may occur multiple times. Some variables are common to all six record types while others are not. For example, the fixed portion of all record types contain demographic information such as beneficiary date of birth and sex, and transaction information such as the date the CWF host site adjudicated the claim. The fixed portion of all institutional records also contains the principal diagnosis and additional transaction information such as payment amounts.

The variables contained in the trailer groups portion of the record vary by record type. For example, all record types contain an edit group, composed of variables that indicate data quality problems. Only physician/supplier and DMERC-processed claims records contain line item variables associated with specific services; many services can be listed on the same claim. All institutional claims contain another (secondary) diagnosis group and a revenue center group.

Institutional data (inpatient/SNF, outpatient, HHA, and hospice) and non-institutional (Physician/Supplier and DME) are available beginning in 1991. DMERC-processed claims record in a special format are available beginning in 1994. (The special format and processing of these claims was phased in through the end of June 1994.)

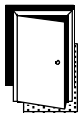
The Division of Data Quality within the Office of Information Services (OIS) serves as the central repository of issues related to HCFA's data. In addition, where there are inaccuracies or shortfalls they must be documented and shared with all users of our data.

It is important that all users report data quality issues to:

Division of Data Quality
HCFA, OIS, SSG, DDQ
N2-14-26
7500 Security Boulevard
Baltimore, MD 21244-1844

E-Mail: DQI@HCFA.GOV
Fax #: (410) 786-0262
Phone #: (410) 786-2860

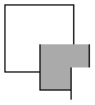
Methods of Access



Access to the NCH 100% Nearline File is available through the Decision Support Access Facility (DSAF) or the 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 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.

DSAF is a menu-driven system for retrieving data from a variety of claims, utilization, and enrollment files including the NCH 100% Nearline File. It can be used for access to NCH 100% state-specific raw claims data, to obtain final action (adjustments resolved) claims data, and for assistance with MANRLINE requests. Users can also obtain state-specific physician/ supplier claims data in final action form. Final action claims data are produced by matching adjustment claims with initial claims and resolving all adjustments. More information on final action data can be found in the introduction to the SAFs.

Migration of Data



Prior to the CWF and the creation of the NCH 100% Nearline File, institutional claims data were stored in five annual bill files: inpatient, HHA, outpatient, hospice, and SNF. Physician/supplier data were stored in files containing payment records. Payment records did not contain line item detail for physician/supplier services. NCH physician/supplier claims records contain detailed information, including submitted and allowed charges and diagnoses, for every line item on the claim.

Since the NCH was implemented in 1991, several versions of NCH record formats have been produced. In 1998, the Nearline was upgraded to a new format (Version 'H') that includes most of the fields coming in from CWF, as well as a number of new derived data fields to enhance the efficiency of the file. Many of the new elements are redundant summary fields that were added to reduce access processing requirements and to simplify logic; e.g., non-institutional line items Allowed Charges, Submitted Charges, and Medicare Payment Amount are accumulated and the totals stored at the claim level. History (back to 1991 service dates) was populated with data in the new fields where derivable from existing information.

The expanded format contains a number of significant data fields, including placeholders for the National Provider Identifier (NPI) and PAYERID fields. Version 'H' also provides for up to two occurrences of Managed Care Organization (MCO) period information, including the MCO Plan Contract Number, which will be a key field for implementing the collection/storage of encounter data pursuant to the Balanced Budget Act. The revised format allows up to five occurrences of Demonstration Identifier information (alias Demo ID; recently renamed Special Processing Number or SPN). The Demo ID/SPN will facilitate the ready identification of demonstration claims in the Nearline File, which is not possible in the current format.

Additionally, fields have been renamed that have caused confusion among users; fields have been resequenced into more logical groupings; field sizes have been increased and filler positions added throughout to accommodate future expansion; and fields that no longer contain any data have been deleted from the record. Two major format changes have also been implemented: (1) all dollar amounts are increased to S9(9)V99 to provide a consistent standard format for all dollar amounts in the Nearline data; and (2) all dates are reformatted to 9(8) to conform to the International Standards Organization (ISO) standard format for millennium compliance and ease of use.

List of Data Elements

Lists of the data elements contained in version H of the NCH 100% Nearline File, along with brief definitions and coding schemes, will be included in the future on the HCFA web site at www.hcfa.gov. The data elements are listed by each of the six version H record types: inpatient/SNF, outpatient, HHA, hospice, physician/supplier, and DME.

STANDARD ANALYTIC FILES

SAFs contain final action claims data in which all adjustments have been resolved. When an error is discovered on a claim, a duplicate claim is submitted indicating that the prior claim was an error. A subsequent claim containing the corrected information may then be submitted. In SAFs, the impact of the error claims and their corrections are netted out.

The SAFs group includes eight files:

- Inpatient 100% SAF
- Outpatient 100% SAF
- HHA 100% SAF
- Hospice 100% SAF
- SNF 100% SAF
- Clinical Laboratory 100% SAF
- DME 100% SAF
- 5% Sample Beneficiary SAF

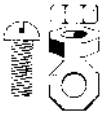
Because the information about file creation, file maintenance, file structure and usage, data structure and usage, methods of access, and migration of data is virtually identical for all SAFs, that information is presented once, followed by descriptions of features unique to each SAF.

File Creation

When analyzing NCH data, researchers must create decision rules and write algorithms to account for claim adjustments. The final action data SAFs were initially created to eliminate the need for this step. Final action data relieve users of the need to account for adjustments and provide a uniform file for analysis purposes.

The SAFs are constructed from weekly data submissions to the NCH 100% Nearline File. Final action claims data are obtained by processing NCH claims through a series of algorithms designed to match original claims with adjustment claims and resolve all adjustments. Files containing the current year's submissions are available beginning with data submitted from January to June of that year. Annual files are created in July for the prior calendar year (based on claim through date); this 18-month window captures approximately 98 percent of the claims for services provided in a given year. Availability of the SAFs is discussed in Data Structure and Usage.

File Maintenance



The SAFs are updated every quarter through June of the following year when the annual files are finalized. Current year's data are created six months into the year; then it is updated quarterly until finalized at 18 months. Once the annual files are finalized, they are not updated to include late arriving claims. Annual files are approximately 98% complete. Because the data for the SAFs come from the NCH 100% Nearline File, they have been subjected to the NCH/CWFMQAS.

File Structure and Usage

Tthe SAFs contain the same variables as the NCH 100% Nearline File records from which they are derived. The files are sorted by Beneficiary Claim Account Number and NCH Category Equatable BIC.

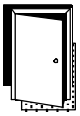
The SAFs can be used to support policy development decisions such as coverage determination and rate setting, quality of care measurement through development of geographic and beneficiary/provider profiles, monitoring of contractor activities, and program research. They are a valuable resource for studies requiring utilization data because adjustments have already been taken into account.

Data Structure and Usage

Each SAF is a different subset of NCH 100% Nearline File data. Thus, the information about record structure, the file sample, and the number of records varies. This information is presented in the individual SAF discussions.

All institutional and non-institutional SAFs (100% & 5%) are available from 1991 onward. Data are stored on tape at the HDC.

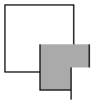
Methods of Access



DSAF is a menu-driven system for retrieving data from a variety of claims, utilization, and enrollment files including SAFs. Using DSAF, the 100% SAFs (inpatient, outpatient, HHA, hospice, SNF, and clinical laboratory) can be subset to obtain records from 1, 5, 20, 60, or 100 percent samples of beneficiaries. In addition, the 5% Sample Beneficiary SAF can be subset to obtain a one percent sample of beneficiaries. DSAF also allows beneficiary cohorts to be defined based on the values of a select group of data elements: beneficiary demographics (e.g., age, sex, and race); record types (e.g, inpatient, outpatient, HHA); data elements unique to the file; and service year.

The SAFs are also available as Beneficiary Encrypted Files (BEFs), with patient, physician, and other identifiers encrypted. The BEFs are described in the Beneficiary Encrypted Files Data chapter.

Migration of Data



Because the SAFs were first created in 1993, they have not been affected by the changing HCFA data processing environment.

List of Data Elements

Data element lists have not been included for the SAFs. However, the data element lists for the corresponding NCH 100% Nearline File record types, which are similar to the SAFs, will be included in the future on the HCFA web site at www.hcfa.gov.

Inpatient 100% Standard Analytic File (SAF)

The Inpatient 100% SAF contains final action inpatient claims data for all Medicare beneficiaries. It has the same structure as the NCH 100% Nearline inpatient/SNF record. However, this file contains only inpatient hospital data; it does not contain SNF data. The record structure is described in Data Structure and Usage for the NCH 100% Nearline File. The 1998 Inpatient 100% SAF contains approximately 12 million records.

Outpatient 100% Standard Analytic File (SAF)

The Outpatient 100% SAF contains final action outpatient claims data for all Medicare beneficiaries. It is available from 1986 to 1988 in one format, and from 1989 onward in another format and has the same structure as the NCH 100% Nearline File outpatient record and are described in Data Structure and Usage for the NCH 100% Nearline File. The 1998 Outpatient 100% SAF contains approximately 105 million records.

Home Health Agency (HHA) 100% Standard Analytic File (SAF)

The HHA 100% SAF contains final action HHA claims data for all Medicare beneficiaries. It has the same structure as the NCH 100% Nearline HHA record. The record structure is described in Data Structure and Usage for the NCH 100% Nearline File. The 1998 HHA 100% SAF contains approximately 12 million records.

Skilled Nursing Facility (SNF) 100% Standard Analytic File (SAF)

The SNF 100% SAF contains final action SNF claims data for all Medicare beneficiaries. It has the same structure as the NCH 100% Nearline inpatient/SNF record. However, this file contains only SNF data; it does not contain inpatient hospital data. The record structure is described in Data Structure and Usage for the NCH 100% Nearline File. The 1998 SNF 100% SAF contains approximately 3.3 million records.

Hospice 100% Standard Analytic File (SAF)

The Hospice 100% SAF contains final action hospice claims data for all Medicare beneficiaries. It has the same structure as the NCH 100% Nearline hospice record. The record structure is described in Data Structure and Usage for the NCH 100% Nearline File. The 1998 Hospice 100% SAF contains approximately 1 million records. This file is the only Medicare claims file that provides hospice data alone.

Clinical Laboratory 100% Standard Analytic File (SAF)

The Clinical Laboratory 100% SAF contains final action claims data for all clinical laboratory services reported on physician/supplier claims. Clinical laboratory services are identified by a predefined laboratory file (containing all lab HCPCS).

This file has the same record structure as the NCH 100% Nearline physician/supplier record. The record structure is described in Data Structure and Usage for the NCH 100% Nearline File. The 1998 Clinical Laboratory 100% SAF contains approximately 129 million records. This file is the only Medicare claims file that provides clinical laboratory data alone.

Durable Medical Equipment (DME) 100% Standard Analytic File (SAF)

Beginning October 1, 1993, HCFA changed the way it handled the reporting of DME claims. The establishment of DMERCs regionalized the processing for most of these claims. The transition period for this new procedure occurred between October 1, 1993 and June 30, 1994. During the phase-in, DME suppliers submitted claims for payment using the old method and the new. As a result, DME claims for 1993 and 1994 could be reported as either a Record Identification Code (RIC) "O" or a RIC "M." The 1998 file contains approximately 41 million RIC "M" and 2 million RIC "O" records. The 1994 DME file (RIC "M") contains claims for the last quarter of 1993 and includes all 1994 DME claims submitted to DMERCs. There will continue to be some DME claims submitted through local carriers as RIC "O."

5% Sample Beneficiary Standard Analytic File (SAF)

The 5% Sample Beneficiary SAF contains all final action claims data for a five percent sample of Medicare beneficiaries. Beneficiaries with a HIC number with the digits 05, 20, 45, 70, or 95 in the

eighth and ninth positions are included in the sample. It has the same structure as the NCH 100% Nearline File. The record structure is described in Data Structure and Usage for the NCH 100% Nearline File.

The 5% Sample Beneficiary SAF includes the following component files: Inpatient, Outpatient, HHA, Hospice, SNF, Physician/Supplier, which includes pre-October 1993 DME data, and, effective March 1997, DME RIC "M" data from October 1993 onward. Prior to March 1997 these files did not include RIC "M" DME data. These 5% files are also available individually. The 1998 5% Sample Beneficiary SAF contains approximately 39 million records. The 5% Sample Beneficiary SAF is the only SAF that contains physician/supplier data. It replaces the BMAD Beneficiary File for accessing physician/supplier claims data. Institutional data are available in this file beginning in 1989; physician/supplier data are available from 1991 onward.

STAY RECORDS FILES

A stay record summarizes all services rendered to a beneficiary from the time of admission to a facility through discharge. The only stay record data currently produced by HCFA are contained in the MEDPAR File.

Medicare Provider Analysis and Review (MEDPAR) File

The MEDPAR File (commonly referred to as MEDPAR) contains inpatient hospital and SNF final action stay records. Each MEDPAR record represents a beneficiary stay in an inpatient hospital (where discharged) or in a SNF (may still be a patient.) Only inpatient records with discharge dates are included in MEDPAR; SNF records are included even if discharge data are not present because discharge information is not always received. These records are a major source of data for program analyses, evaluations, and utilization statistics.

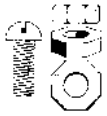
File Creation

CWF claims records are processed weekly into the NCH Nearline Repository. Beginning in June 1995, the Inpatient and SNF claims from the NCH 100% Nearline File became the source of MEDPAR. (Prior to June 1995, MEDPAR was created from claims from the Medicare Quality Assurance System). Each month Inpatient and SNF claims are accumulated from the Nearline Repository. At the end of each quarter the monthly files are merged into a database that contains claims for the current and previous two years. The updated database is processed through the final action algorithms, and then split into two segments for each year. One segment contains all Inpatient claims with discharge dates and SNF claims with admission dates in January through September.

The other segment contains claims with dates in October through December. This allows for the creation of fiscal year and calendar year files as needed.

Remaining claims are summarized to create a stay record. A match on claim number, admission date, and provider number creates a stay. These stay records are then used to create MEDPAR records.

File Maintenance



The MEDPAR files are created quarterly in March, June, September, and December. The claims records used to create MEDPAR are validated paid claims that have been processed by the NCH Medicare Quality Assurance System. No additional edits or validation are performed before the MEDPAR records are created. However, OIS QA initiatives have been directed at MEDPAR data to uncover any specific reliability and consistency problems with the data. Any data anomalies are addressed through warning indicators contained in the record. These indicators flag questionable data in each MEDPAR record.

From 1979 through 1983, MEDPAR contained stay data for a 20% sample of Medicare beneficiaries. These were calendar year files only. In 1984, MEDPAR was expanded to contain stay data for all beneficiaries. Annual MEDPAR files, identified by the file update date, are available for fiscal years and calendar years 1984 forward.

File Structure and Usage

MEDPAR records contain data from claims for services provided to Medicare beneficiaries admitted to Medicare-certified hospitals and SNFs. Each MEDPAR record may represent one claim or multiple claims, depending on the length of a beneficiary's stay and the amount of inpatient services used throughout the stay. Approximately 95% of inpatient stays and 50 % of SNF stays involve a single claim.

Both beneficiary HIC number and provider number are included on each record, making MEDPAR a valuable data resource for health care research and analyses that require beneficiary or facility-specific data. The database is sorted by HIC number, admission date, and provider number (all in ascending order) to enable an individual's stay records to be linked to study inpatient history and track patterns and outcomes of care over time. Beneficiary demographic characteristics, diagnosis and surgery information, and use of hospital or SNF resources are included on the MEDPAR records. In addition, the records contain detailed accommodation and departmental charge data, and days of care and entitlement data.

As of June 1995, MEDPAR records represent final action claims data in which all adjustments have been resolved. Prior to June 1995, a MEDPAR record represented an accumulation of adjustment claims. Therefore, the number of bills used to make up a stay will generally be one beginning with the June 1995 update, whereas it would have been several claims prior to June 1995.

MEDPAR has been used for research into chronic diseases prevalent in the elderly population such as cancer, heart disease, and diabetes. It has also been used to update annual hospital Prospective Payment System (PPS) rates.

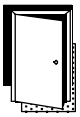
Data Structure and Usage

Since it was first developed, MEDPAR has been enhanced to accommodate the increasing level of detail required on the claims submitted by providers. The claims, and in turn MEDPAR, include detail regarding accommodations, services, and costs associated with each inpatient stay at a hospital or SNF. The inclusion of this information was critical for the implementation of PPS, and also supplied policy analysts and researchers with data to track patterns and outcomes of care provided to Medicare beneficiaries.

MEDPAR is available in an enhanced 500 character, fixed-length record format for fiscal and calendar years 1987 to the present. As of March 1999, a 514 character, fixed-length record format was created to accommodate Y2K date field expansions. In March 2000, the 500 character MEDPAR will no longer be available. Fiscal and calendar years 1984 to 1986 are available in abbreviated formats. Accommodation data and service data fields have been enhanced to provide additional detail on specific types of hospital resources used (e.g., whether a stay was in a private or semi-private room in an intensive care or coronary care ward; or whether coronary care was provided for myocardial infarction or pulmonary care). In addition, new fields contain warning indicators which flag claims with questionable data, such as multiple claims with missing interim claims, a zero or negative payment or total charges, or a date of death earlier than date of admission.

The 1997 MEDPAR File contains approximately 14 million records: 12 million hospital records and 2 million SNF records.

Methods of Access



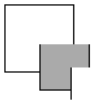
Access to MEDPAR is restricted because it contains patient identifiers. However, three BEFs with patient identifiers are maintained:

- the Expanded Modified MEDPAR-Hospital File (National)

- the Expanded Modified MEDPAR Hospital File (State - a subset of the national data extracted by provider state or beneficiary state of residence)
- the Expanded Modified MEDPAR-SNF File

Further information about these BEFs can be found in the Beneficiary Encrypted Files Data chapter.

Migration of Data



Prior to the CWF, MEDPAR was created from bill records. Since the adoption of the CWF, MEDPAR is created from CWF claims records processed through the NCH Medicare Quality Assurance System. As of June 1995, MEDPAR is created from the NCH 100% Nearline File inpatient/SNF claims records processed through final action algorithms.

List of Data Elements

The list of data elements to be included on the HCFA web site (www.hcfa.gov) for MEDPAR will reflect the current, expanded version of the full MEDPAR File in its 500 character record format (6/95 format.) MEDPAR files created in earlier years contain less detailed information and a smaller number of variables per record.

PART B MEDICARE FILES

The Part B Medicare Files contain data on the utilization of physician services and certain other medical services, equipment, and supplies covered by the Medicare SMI plan. Prior to the implementation of the CWF, physician and supplier claims data processed by carriers were submitted to HCFA in the form of payment records. Payment records, however, did not include line item detail such as payment amount for each service. To supplement its physician/supplier utilization data, HCFA required all carriers to submit four summary and sample files extracted from the complete history of the Medicare claims they processed. These four historic BMAD files included the following:

- BMAD Procedure File
- BMAD Prevailing Charge/Pricing File
- BMAD Provider File
- BMAD Beneficiary File

The BMAD files were derived from only those Medicare claims processed by Medicare carriers. Claims for home health services and outpatient hospital services, while also Part B benefit claims, were processed by FIs and hence were not included in the BMAD files. Physician/supplier services

rendered to beneficiaries enrolled in Health Maintenance Organizations (HMOs) were also not included.

After 1990, annual carrier BMAD submissions were discontinued because complete physician/supplier claims data became available on a flow basis from the CWF. In addition, the BMAD Prevailing Charge/Pricing File became outdated with the implementation of the physician fee schedule, and fee schedules for clinical laboratory services and DME. As a result, the BMAD files have been superseded by files created from other data sources.

The following table lists the historic BMAD files and their successors:

BMAD File	**Successor File(s)**
BMAD Procedure File	! Physician/Supplier Procedure Summary File
BMAD Prevailing Charge/Pricing File	! DMEPOS (See PUF Chapter) ! Annual Physician Fee Schedule Transition Payment Amount File (See PUF Chapter) ! Clinical Diagnostic Lab Fee Schedule File (See PUF Chapter)
BMAD Beneficiary File	! 5% Beneficiary Physician/Supplier data (subset of 5% Sample Beneficiary SAF)
BMAD Provider File	! Physician Sample File

The successors to the BMAD Procedure File and the BMAD Beneficiary File are created from CWF claims data. The successors to the BMAD Prevailing Charge/Pricing File are fee schedules containing pricing amounts for individual services. The successor to the BMAD Provider File, the Physician Sample File, is created from the NCH file.

Physician/Supplier Procedure Summary (PSPS) File

The PSPS Procedure File provides information on services used by Medicare beneficiaries during a calendar year. The information is provided at the procedure level.

File Creation

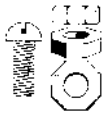
Each year, HCFA produces the PSPS file using all the claims for physician and supplier services rendered to Medicare beneficiaries during the year and processed by the carrier through June 30 of the following year. Each carrier summarized the claims data (i.e., submitted and allowed charges)

to the procedure code level. Every procedure code reported on a claim in a given year was represented in the PSPS File even if it occurred only once.

HCFA receives submissions from carriers via CWF and creates a PSPS File once a year, generally six months after the end of the service year. HCFA processes the files through a series of edit and validation programs. These programs included checks for valid codes and logical consistency between related data elements within each record. Records identified in the edit process as being incorrect but within acceptable tolerance levels are annotated as errant records and included in the file.

Since 1991, HCFA has produced a Physician/Supplier Procedure Summary File from CWF claims data as a successor to the BMAD Procedure File.

File Maintenance



The PSPS is created on an annual basis. It is available as of July of the succeeding year. For example, 1998 incurred year is available in July of 1999.

File Structure and Usage

The PSPS File contains frequency and charge data for every procedure occurring on the Medicare claims processed by all carriers. Each PSPS File record represents a unique combination of the following:

- Carrier number
- HCPCS procedure code
- First modifier
- Second modifier
- Type of service
- Place of service
- Specialty
- Locality

The PSPS File provides information on the services used by Medicare beneficiaries at the local and national level, and the cost of these services to the SMI fund. Such information is crucial for the evaluation of the impact of alternative Medicare fee schedules and payment policies, and for Medicare utilization trend analyses and projections.

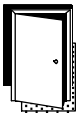
Data Structure and Usage

The PSPS File contains records representing unique combinations of carrier number, HCPCS code, first and second modifiers, type of service, place of service, specialty, and locality. The file includes one record for each combination occurring at least once in the combined claims of the reporting carriers. The more recent annual PSPS files contain approximately nine million records.

PSPS File records contain each of the data elements used to define a unique record (e.g., carrier number and type of service), as well as procedure frequency, and total charges submitted, allowed, and denied. Also included are the miles, time, units, or services for each procedure depending on the type of procedure. Records are sorted by locality, specialty code, type of service, place of service, and procedure and modifier codes.

Known data limitations at the individual data element level are described in detail in the Center for Health Plans and Providers (CHPP)-produced *PSPS File Limitations* document. This document includes year- and carrier-specific examples of errant records, and contains tables illustrating the amount and percentage of total allowed charges contained in errant records.

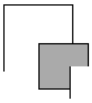
Methods of Access



A public use version of the PSPS File, described in the Public Use Files (PUFs) Data chapter, is available for purchase.

Data from the Physician/Supplier Procedure Summary File (created from CWF claims data), are available online at the HDC through the Part B Extract and Summary System (BESS). BESS provides three complete years of data plus the current year to date. This interactive system is designed to access and extract select physician/supplier summarized claims data and generate pre-formatted reports for the current and three prior years. BESS is open to all HCFA HDC users.

Migration of Data



Since 1991, HCFA has produced a Physician/Supplier Procedure Summary File using CWF claims data. The Physician/Supplier Procedure Summary File differs from the BMAD Procedure File in that it captures CWF data quarterly, is maintained and made available on a year-to-date basis, and contains the Medicare payment amount. The calendar year file is closed six months after the end of the service year (instead of the three-month lag allowed for the historic BMAD files). Data for the current calendar year and the three previous years are available through BESS.

Since the Physician/Supplier Procedure Summary File is created from daily submissions of CWF data, it is more current than the BMAD Procedure File. Year-to-date files are available one month after the close of each quarter. Full calendar year data are available by July of the following year.

List of Data Elements

A list of the data elements contained in the Physician/Supplier Procedure Summary File, along with brief definitions and coding schemes, will be included in the future on the HCFA web site at www.hcfa.gov.

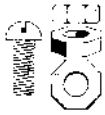
Physician Sample File

The Physician Sample File contains final action claim line item-level data for services rendered to Medicare beneficiaries by a sample of physicians.

File Creation

The Unique Physician Identification Number (UPIN) Validation File is statistically analyzed for the random selection of the physicians and is based on the ending two digits of the UPIN in the Carrier Identification Number and Locality Code. The identified sample is extracted from the NCH file six months after the close of the calendar year. The Physician Sample File is created using all the physician line items extracted for the year and netted into final action. Because the data for this file come from the NCH, they have been subjected to the NCH Quality Assurance System (NCHQAS).

File Maintenance



The Physician Sample File is created eight to 12 months after the close of the calendar year.

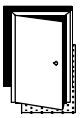
File Structure and Usage

The Physician Sample File contains service information such as place and type of service, and charge and payment amounts. The files contain records for every service furnished to Medicare beneficiaries by a sample of physicians in a given year. Each Physician Sample File record contains data from one physician claim line item. The file is used for research.

Data Structure and Usage

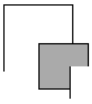
The Physician Sample File is a sequential flat file that resides on robotic cartridge containing fixed-length records sorted in ascending order by HIC Number and Carrier Control Number. The file contains approximately 20 million records per year. The data are derived from all physician line items.

Methods of Access



The Physician Sample File is also available as a BEF and is described in the Beneficiary Encrypted Files Data chapter.

Migration of Data



With the implementation of the CWF, physician claims data became available to HCFA on a flow basis. Annual carrier BMAD submissions and creation of the BMAD Provider File were discontinued after HCFA received the 1990 data. The Physician Sample File was first created in 1991 and is the successor to the BMAD Provider File

List of Data Elements

A list of data elements contained in the file as well as brief definitions and coding schemes will be included in the future on the HCFA web site at www.hcfa.gov.

OTHER UTILIZATION FILES

The Other Utilization Files group contains two databases or files:

- Provider Summary File
- Physician Summary File

The Other Utilization Files pre-date the advent of the CWF (as do the MEDPAR File and BMAD files).

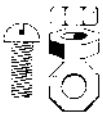
Provider Summary File

The Provider Summary File contains summarized utilization information for Medicare-certified hospitals. The file includes information such as number of claims, days of care, charges, and payment amounts.

File Creation

The Provider Summary File is created from CWF institutional claims records. The claims records received by HCFA are assembled into files each week. Each month, four or five weekly files are merged into a monthly claims file. At the end of each month, the files are summarized by provider. Monthly files are available from 1987 onward.

File Maintenance



Monthly files containing summary data for each provider are processed as described in File Creation.

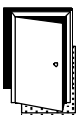
File Structure and Usage

The Provider Summary File contains utilization information summarized by Medicare-certified hospitals, including number of claims, days of care, charges, and payment amounts. The unit of analysis is the provider. HCFA uses this file primarily for monitoring and projecting patterns of hospital admission and length of stay.

Data Structure and Usage

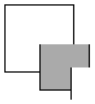
The Provider Summary File is a sequential flat file containing fixed length records sorted by provider number. Each record consists of approximately 40 data elements including the number of discharges and days of care for beneficiaries who are aged and beneficiaries with disabilities, and payment and charge information for each claim type. Monthly files are available beginning in 1987. The July 1995 file contains approximately 309,000 records.

Methods of Access



The Provider Summary File is stored on tape at the HDC.

Migration of Data



The Provider Summary File has not changed significantly with the advent of the current HCFA claims processing environment.

List of Data Elements

A list of the data elements contained in the Provider Summary File, along with brief definitions and coding schemes, will be included in the future on the HCFA web site at www.hcfa.gov.

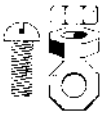
Physician Summary File

The Physician Summary File contains summary information on the total annual Medicare charges, payments, and services (denied/allowed) for participating providers, non-participating providers who accept assignment, and non-participating providers who do not accept assignment. The summaries are based on Specialty Code, UPIN, Carrier Identification Number, and Locality Code.

File Creation

The Physician Summary File is created using all physician line items received for services performed in a given year. HCFA creates this file by extracting information weekly from the Common Working File (CWF) claim records. Weekly extractions are summarized into quarterly final action files based on the last date of service. The annual file is constructed by merging the quarterly files. Because the data for this file comes from the CWF, HCFA subjects the data to the National Claims History Quality Assurance System (NCHQAS).

File Maintenance



The Physician Summary File is created on an annual basis.

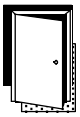
File Structure and Usage

The Physician Summary File summarizes Medicare charges, payments, and services (allowed and denied) for a given service year. The summary keys are Specialty Code, UPIN, Carrier Identification Number, and Locality Code. The Provider categories are totaled based on participating providers, non-participating providers who accept assignment, and non-participating providers who do not accept assignment. The file is used for research purposes.

Data Structure and Usage

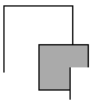
The Physician Summary File is a sequential flat file that contains fixed-length records sorted by Specialty Code, UPIN, Carrier Identification Number, and Locality Code. The file contains approximately 800,000 records per year; the data are derived from all physician line items.

Methods of Access



The Physician Summary File is available for research and other approved program purposes. The file is stored on robotic cartridges at the HCFA Data Center.

Migration of Data



The Physician Summary File is available from 1991 to the present.

List of Data Elements

A list of the data elements contained in the Physician Summary File, along with brief definitions and coding schemes, will be provided in the future on the HCFA web site at www.hcfa.gov.