Appendix E.—Medicare Part A Data Systems

Introduction

The Medicare Part A (Hospital Insurance) data systems were developed and designed for their use in the administration of the Medicare program. Hence, the administrative requirements of the past have governed their content, quality, and timeliness. The Part A data systems provide a record of reimbursement-related actions and the information on which those actions were taken. Although they also provide an information base for program and policy analyses and research, these functions have been secondary. Public use tapes exist for only two of the files discussed in this appendix, a situation that limits the ability of outside researchers to pursue independent research. Nevertheless, the Part A data systems provide a rich base, often the only database, for evaluating some of the important impacts of Medicare's prospective payment system (PPS).

Basic Data Files

Medicare Part A data systems contain data in four major categories:

- *Beneficiary enrollment and eligibility status*—the characteristics of and benefits available to the Medicare enrolled population. The records containing these data elements originate in the Social Security Administration (SSA).
- *Inpatient hospital* claims—characteristics of the patient, the services rendered, and the charges billed during a hospital stay. These data originate in hospital bills.
- *Hospital* costs—capital and operating costs of Medicare-certified hospitals. These data are submitted each year directly by hospitals to intermediaries on Medicare cost reports.
- *Provider* characteristics—attributes of the institutions and kinds of services they provide to Medicare patients. Data on hospital characteristics are available from both the Medicare cost reports and the process of certifying hospitals for treatment of Medicare patients.

The data elements in these categories arrive at the Health Care Financing Administration (HCFA) through a variety of channels and are processed into specific files[1] including:

- the Health Insurance Master Enrollment Record File (referred to as the HIM File), which contains up-to-date information on each Medicare beneficiary;
- the Hospital Stay Record File, which contains a record of each inpatient hospital stay;
- the Hospital Cost Report Information System (HCRIS), which contains data extracted from the Medicare cost reports; and
- the Provider of Services (POS) File, which contains data items pertaining to the Medicare hospital certification process.

These files, which can be further manipulated or merged as the need arises, are the basic sources for all derivative files that may be created either to support the operations of the health insurance programs or to monitor and evaluate the performance or impact of the system. Each of these data files is described in greater detail below.

Health Insurance Master Enrollment Record (HIM) File

The HIM File contains information supplied to HCFA by SSA. It includes basic identifying and demographic information about Medicare Part A beneficiaries and also indicates whether the beneficiary is covered under Part B of Medicare (Supplementary Medical Insurance), the extent to which certain limited benefits are still available to the beneficiary, and the beneficiary's current status regarding deductible and coinsurance.

The HIM File is updated frequently to reflect changes in beneficiaries' eligibility for benefits. All hospital claims arriving at HCFA are passed against the appropriate beneficiary's record within a few days, and in that process, some information about the beneficiary is also appended to claims records for use in the claims files. The HIM File also contains current information on Medicare inpatient admissions, because with each admission the hospital makes a "query" to HCFA through the fiscal intermediary to confirm the patient's eligibility for Medicare coverage. This query is answered by checking the HIM File, and in that process, the beneficiary's record is flagged as a current hospitalization.

Hospital Stay Record File

Prior to October 1983, the Hospital Stay Record File aggregated individual bills pertaining to a hospital stay for a 20-percent sample of Medicare beneficiaries (340). Diagnostic and surgical information were coded (using the ICD-9-CM[2] classification) from narratives for this 20-percent sample.

[1] For the most part, these files are automated and can thus be considered to exist on computer tapes or disks.

[2] ICD-9-CM, the International Classification of Diseases, 9th Revision, Clinical Modification, is designed for the classification of morbidity and mortality information for statistical purposes, and for the indexing of hospital records by disease and operations, for data storage and retrieval (322).

In October 1983, hospitals were required to provide diagnostic and procedural data for all Medicare bills, and under PPS, the hospital will submit only one bill per hospital stay. Since October 1983, the Hospital Stay Record File has included diagnostic and procedural information as well as provider and beneficiary demographic data for 100 percent of Medicare hospital stays. Data on diagnoses, DRG assignment, enrollee demographic data, major procedures performed, and patient charges by department are taken from the HIM File and patient billing data supplied by the fiscal intermediaries (contractors authorized by HCFA to make Part A payments to hospitals). The file is updated quarterly.

The procedures for reporting diagnostic information included in the Hospital Stay Record have changed substantially since these data became a reporting requirement. Beginning in 1977, fiscal intermediaries were required to report the first 45 characters of diagnostic information and the first 41 characters of surgical information reported on the provider bill for the 20-percent sample on which diagnostic information was reported. HCFA used an automated system to code the narratives. When codes could not be established using the automated process, they were coded manually.

In 1981, HCFA began to allow institutional providers to report coded ICD-9-CM information rather than narrative. In October 1983, the reporting of this information using ICD-9-CM cases became a requirement rather than an option. Also in October 1983, HCFA began to require that hospitals report up to four additional diagnostic and two additional surgical procedure codes in addition to the principal diagnostic and surgical codes that had been previously required.[3]

The currency of the Hospital Stay Record File depends largely on the timeliness of bill submission by hospitals and processing HCFA and SSA. The vast majority of bills are submitted within 1 month of discharge, but some bills are submitted up to 18 months later. The average time is about 1 month, and one-half of all bills are submitted within 17 days (394). Intermediaries take between a week and 10 days to put bill data on magnetic tape, edit the data for consistency and completeness, and send them to HCFA. Thus, at least 50 percent of the bills are processed within 2 months, but a small proportion are subject to delays that are much longer. Transit and batching time at HCFA takes an average of 2 to 3 weeks (394).

[3]As a consequence of these changes, Medicare has adopted a new billing form, the UB-82, which provides room for the additional diagnostic and procedural data as well as more detailed charge data by revenue-producing cost centers. The UB82 is currently being phased in and is expected to be universal by the end of this year (119).

Hospital Cost Report Information System (HCRIS)

HCRIS is a new automated database of selected information extracted from Medicare cost reports. The cost report is submitted annually by each hospital within 4 months of the end of its fiscal year to the fiscal intermediary. Until the advent of PPS, the Medicare cost report was the primary document on which hospital payment was based.[4]

The cost report form has changed frequently and become more complex in response to changes in law and regulations pertaining to hospital reimbursement. For the most part, the changes in the cost report form have represented the addition of more detailed data or reorganization of existing data, Several important kinds of data have been consistently available over the lifetime of the Medicare cost reports.[5]

The cost report consists of a number of worksheets (analogous to income tax forms) which require the hospital both to provide statistical and financial data and to perform calculations to arrive at a level of reimbursement. The data items included in HCRIS are those identified by HCFA as necessary to meet the most often used and highest priority data needs. The Medicare cost reports include the following basic elements:

- hospital statistics—selected characteristics of the hospital, including:
  - —type of ownership or control,
  - —number of beds available,
  - —inpatient days,
  - —average number of employees on payroll,
  - —number of admissions;
- operating expenses by hospital cost center;[6]

[4]The cost report is still necessary for computing hospitals' payments both because of the 3-year transition period and the exclusion of certain costs (e. g., outpatient, capital, medical education) from DRG payment. HCFA is required to keep hospital cost reports until Sept. 30, 1988 (Public Law 98-21), but the form of those reports may change.

[5]The most recent cost report form, HCFA-2552-84, is to be used by hospitals reporting on fiscal years beginning on or after Oct. 1, 1983. This form, reporting on the first year of the PPS, is as extensive as its predecessors. It consists of a total of 112 pages of forms and worksheets, although not all of the data items are applicable to any given hospital. The changes from the previous form (HCFA-2552-83), e.g., items pertaining to paramedical expenditures, appear to be relatively minor additions as required by PPS. HCFA'S latest proposed form, HCFA-2552-85, for fiscal years beginning on or after oct. 1, 1983, is currently being circulated outside of HCFA for comments (117).

[6]While specific cost centers are defined on the cost report, hospitals, may, with the approval of the intermediary, combine or break down additional cost centers as they desire. Hospital cost centers include revenue-producing units such as radiology or pharmacy as well as non-revenue-producing units such as housekeeping or laundry. They also include outpatient cost centers as separate entities.

- depreciation expenses by type (e.g., movable equipment, buildings and fixtures);
- capital expenditure data by project;
- total charges by cost center;
- Medicare charges by cost center;
- routine inpatient nursing salary costs;
- malpractice costs; and
- standard financial statement information—assets, liabilities, income, etc.

Hospitals must submit a cost report to their fiscal intermediary within 3 months of the close of their fiscal year; they routinely receive a 30-day extension. The fiscal intermediary then has up to 1 year to review, audit, and finally settle on the amount due to the hospital. In this process, some data items on the cost report as submitted may be challenged and changed by the fiscal intermediary. Thus, the final settled cost report data present a more accurate picture of allowed hospital costs than do the submitted reports, but they also involve a substantial lag time (one additional year). The difference between the "as submitted" and "as settled" cost reports for fiscal year 1982 was estimated to be $700 million, or about 2 percent of total Medicare hospital payments (72).

On paper, HCRIS is a powerful addition to the Medicare data systems. Prior to its implementation in 1982, HCFA had no automated files for Medicare cost reports. When cost reports were required for some aspect of program operation or analysis, hard copies would have to be obtained from the intermediaries. For example, in computing the first weights for diagnosis-related groups, HCFA actually entered data from the most recent copy of each hospital's report which were obtained directly from the fiscal intermediaries (229). There is also a shortened public use version of this data file, the HCRIS-180.

As of June 1985, the HCRIS file of cost reports for hospitals' fiscal years ending on or before January 1, 1982, was reasonably complete and consisted primarily of "as submitted" reports. The file for fiscal years ending between January 2, 1982 and September 29, 1983, was about 80 percent complete and primarily consisted of settled reports. The file of cost reports for fiscal years ending between September 30, 1983 and September 29, 1984, data reflecting the first year of regulations under the Tax Equity and Fiscal Responsibility Act of 1982 (Public Law 97-248), was about 15 percent complete and primarily consisted of settled cost reports (377).

About 90 percent of the data in the HCRIS database is received by HCFA in automated form from fiscal intermediaries; the remainder (largely from those hospitals with low Medicare utilization or an all inelusive/no charge structure) is received in hard copy.

Although HCRIS is intended to contain data from the cost reports submitted by hospitals prior to settlement as well as the final settled cost reports, HCFA has temporarily suspended the transmission of "as submitted" cost reports, purportedly because of the fiscal intermediaries' workload problems (340). HCFA's Bureau of Data Management and Strategy has recommended the transmission of the "as submitted" cost reports for the first year of PPS, hospitals' fiscal years ending between September 30, 1983, and September 29, 1984.[7] If this recommendation is approved, the "as submitted" file for fiscal year 1984 could be about 80 percent complete by April 1986 (377).

The expected timeframe for the completion of the HCRIS file for final settled cost reports is at least 17 months from the end of the hospital's fiscal year: 4 months for hospitals to submit, 12 months for the fiscal intermediaries to settle, and 1 month for the fiscal intermediaries to transmit (340). Since the file for any given fiscal year includes 200 to 300 hospitals whose fiscal year ends August 31 (11 months after the September 30 fiscal year end for many other hospitals), the whole reporting cycle takes approximately 28 months if these hospitals are included in the file for the same fiscal year (170). To this time lag must be added any delays in the issuance of the reporting forms, as is the case currently. Of course, this time lag could be shortened by the utilization of "as submitted" cost reports, but this involves some loss of accuracy and, in any event, there is no "as submitted" cost report data on the HCRIS file at present.

Provider of Services (POS) File

The POS File contains information on approximately 7,000 hospitals, 5,200 skilled nursing facilities (SNFs), 3,000 home health agencies, 3,500 independent laboratories, and 1,600 other Medicare participating facilities. Data from Medicare certification applications and surveys form the basis for this file, which is updated on a daily basis with new certification/recertification information (325). Recertification information replaces preexisting information, which is then placed in a history file.

The POS record varies with each kind of facility and includes some data specific to the particular provider type. Among the data included are:

- facility location (i. e., city, State, county, Stand-

[7]HCFA-2552-84 was recently approved by HCFA and distributed to the fiscal intermediaries and hospitals by mid-April and the end of April, respectively. Fiscal year 1984 cost reports are due within 60 days after the forms and instructions are received. Obviously, this will cause additional delay for the completion of the HCRIS file for the first year of PPS.

ard Metropolitan Statistical Area, HCFA region, responsible professional standards review organizations);
- **facility type;**
- **type of control;**
- **total number of beds;**
- **number of certified beds;**
- **services offered; and**
- **number of registered nurses, licensed practical** nurses, pharmacists, social workers, occupational therapists, speech therapists, physical therapists, and other employees.

The data for each facility in this file are identified by the same provider number used in billing and claims records, making it possible to link these data to the provider records.

Derivative Files

The basic files described above are large, either because they contain many records, many data entries on each record, or both. Derivative files are often created to extract a smaller number of records or selected data items, thus reducing the data processing burden. Derivative files are also often created when data from two or more basic files need to be linked together for the sake of analysis. Medicare's Part A data systems contain three such derivative files:
- the Medicare Provider Analysis and Review (MEDPAR) File;
- the Medicare History Sample File; and
- the Medicare Automated Data Retrieval System (MADRS).

Each of these files is described below,

Medicare Provider Analysis and Review (MEDPAR) File

Records in the MEDPAR File are created by merging data on the hospital's characteristics and costs with a shortened version of the Hospital Stay Record File. The MEDPAR records include the data on diagnoses, procedures, and departmental charges that are contained in the Hospital Stay Record. Although the Hospital Stay Record File since October **1983** has included data on 100 percent of hospital stays, MEDPAR will continue to report on a 20-percent sample of beneficiaries. The **1984** MEDPAR File, like the Hospital Stay Record File for that year, will record up to four diagnoses and two procedure codes in addition to the principal diagnostic and surgical codes. There is also a public use version of this file.

The MEDPAR File is created on a quarterly basis. The most current full year is the 1983 MEDPAR File. In June 1985, HCFA was still working on the completion of the fiscal year 1984 MEDPAR File, and the file was 82 to 83 percent complete. Normally, the fiscal year 1985 MEDPAR File would have been expected to be 94 to 95 percent complete by April 1985. The delay is apparently due to the transition to the PPS and data processing problems by the fiscal intermediaries caused by the transition to UB-82 (119). In the future, HCFA expects the file for any fiscal year will be available within 3 to 6 months of the end of that fiscal year. Because the time taken to submit bills is to a large extent outside of HCFA's control, it is expected that the MEDPAR file will be only 90 to 95 percent complete within this timeframe.

Medicare History Sample (MHS) File

The Medicare History Sample File is based on a .5-percent sample of Medicare enrollees for 1974 and later and provides a history of the utilization of Medicare services for enrollees included in the sample. Selected enrollee characteristics are obtained from active and inactive HIM files. Utilization data for each year are obtained from inpatient hospital, SNF, home health agency, and physician office and outpatient bills processed by Medicare intermediaries and carriers.

A single record for each Medicare beneficiary is updated on an ongoing basis with utilization data from the billing records as well as characteristics data from the HIM records. Once a beneficiary is included in the Medicare History Sample, that person remains on the file regardless of utilization activity or death. A 5-percent sample of new Medicare enrollees is added each year,

One section of the Medicare History Sample record identifies the demographic characteristics of the enrollee, the basis for entitlement, and where applicable, date and reason for termination from Parts A and B of the program. This section is created when the enrollee is first added to the Medicare History Sample file.

Annual additions to the demographic records include data that may change over time, such as whether the enrollee is entitled under Part B and/or Part A, whether coverage under Part B has been accepted, the State, county, and zip code of residence, and whether the enrollee is a participant in a group prepaid practice plan or health maintenance organization.

The remaining sections of the Medicare History Sample record contain charge, reimbursement, and utilization data for outpatient services supplied by an institutional provider, hospital inpatient stays, home health agency visits, SNF stays, physician services, and other services reimbursable under Medicare. The in-

patient stay section of the Medicare History Sample provides information on principal diagnosis and surgical procedure. It is expected that the file will contain the expanded diagnosis and procedure information obtainable from the *new* Hospital Stay Record File (119).

HCFA's Bureau of Data Management and Strategy, which is responsible for developing and maintaining the Medicare History Sample file, is currently working on completion of the **1982** file. The Bureau anticipates that in the future the Medicare History Sample file will be completed 12 months after the end of the calendar year. The 1982 data, however, are taking at least twice the amount of time to complete. The latest available Medicare History Sample file is for 1981 and contains 1,900,000 records, an increase of approximately 100,000 records over the previous year (261).

Medicare Automated Data Retrieval System (MADRS)

Currently under development, the MADRS data file is intended to reorganize and merge Medicare's Part A and Part B claims files to facilitate research and make analysis less expensive. When it is completed, MADRS will sort the claims records in HCFA's files first by the year, then by the county in which the beneficiary resides, and finally by the beneficiary's health insurance number.

This presorted file will enable researchers to access the full array of Medicare claims made by all kinds of providers for a set of beneficiaries without sorting through an entire year's worth of Part A and Part B claims files.[8] MADRS will also contain an index to indicate the location of records for a particular county, beneficiary, or provider (other than individual physicians) on the automated data file, making access to particular records more convenient.

Because of a number of contract difficulties and resource limitations, as of May 1985 the file does not exist, and there is no indication of when it will be completed, although HCFA personnel indicate that there is a firm intention to do so.

[8]HCFA estimates that the cost of sorting through 1 year of the claims files to be about $15,000 (329).