

National And Regional Databases: The Big Picture

Regina McPhillips, DrPH, RN, FAAN

The key to measuring and understanding what effective health care is and to ensuring a consensus among patient, payers, and providers on what care represents the best value, is the availability of comprehensive data and information on the effects of health care choices on patient well-being relative to the costs of treatment. Data is critical to the evaluation of trends in health care delivery and in identifying and promoting health care interventions that are of consistent high quality, appropriate, and cost effective.

As the principal Federal agency responsible for paying for the health care of nearly 50 million Americans, the Health Care Financing Administration (HCFA) routinely collects a great deal of information on the service use and costs of care for its beneficiaries and recipients. The paper will describe these databases. HCFA's program data on its beneficiaries already comprise the largest population, provider and claims databases in the country. HCFA has recently restructured its claims data systems to provide a more efficient mechanism for claims processing, known as the Common Working File, and a single large database, called the National Claims History, to be used for monitoring HCFA programs and research on the effectiveness of care. The national Claims History Database contains information about all Medicare services (except those provided in HMOs). Researchers will be

able to use the National Claims History for information about the care provided in all types of settings, and about adverse patient outcomes.

A summary of the current activities, and possible future directions to expand data availability, improve its quality and usefulness, and focus research on effectiveness of care, will be provided. Emphasis will be placed on the critical need for better information in improving outcomes and enhancing quality of care, and on the importance of valid, complete, and accurate data in identifying and promoting health care interventions that are appropriate and cost effective.

We are all aware that the health care system is characterized by complex interactions among consumers, providers, and payers. It is critical that we be able to measure and understand what effective health care is in its American context and then work to develop a consensus among patients, providers, and payers as to what care represents the best overall value. The key to such policy formulation is the availability of comprehensive data and information on the effects of health care choices on patient well-being relative to the costs of treatment. Indeed, it is difficult to imagine any potentially successful new interventions in the health care system which are not data-driven.

HCFA DATA

HCFA, in carrying out its responsibility to manage the Medicare and Medicaid programs and related quality assurance programs, manages the largest health care database in the world. These data are critical to the evaluation of trends in health care delivery and in identifying and promoting health care interventions that are of consistent high quality, appropriate, and cost effective. HCFA program data play a key and growing role in the pursuit of health care system efficiency. Accurate HCFA program data are essential to the realization of strategic Departmental goals, notably the second of the Department's nine current program directions—to expand the use of cost effective health care practices to ensure the quality and affordability of health care.

As part of HCFA's mission, data have been collected for administrative purposes since the start of the Medicare and Medicaid programs in 1966. Although HCFA had occasionally provided data for epidemiological research in the past, only recently did we begin to make our data available to the entire community. Given that the Department of Health and Human Services alone spent nearly \$150 billion in CY 1990 on health care, it is imperative that the Department place a high priority on assuring that its beneficiaries receive appropriate and cost-effective care, while constantly searching for ways to reduce the administrative costs of its programs. Beyond this, DHHS has extensive responsibilities for improving biomedical research and assuring the dissemination of relevant data to the wider health care community which consumes nearly 12 percent of the GNP.

SUMMARY OF OLD ENVIRONMENT

HCFA's program data on its 34 million beneficiaries already comprise the largest population, provider and claims databases in the country. In the past, Medicare contrac-

tors were required to contact HCFA to confirm eligibility for program benefits. Essentially, upon receipt of a Medicare claim, the contractor forwarded a query to HCFA's Bureau of Data Management and Strategy (BDMS)—the Bureau of which I am the Director. BDMS' system confirmed entitlement and forwarded a response to the contractor. Queries were required on all Part A claims. In addition, contractors queried on Part B claims beginning with the first claim they received each year and continuing until notice that the deductible had been met was received by the contractor.

Although query histories existed, they did not contain comprehensive or accurate information related to Medicare payment. Instead, contractors were required to forward multiple data sets to HCFA. For example, UNIBILLS were forwarded for all services processed by intermediaries. Carriers created and transmitted payment records for noninstitutional Part B services where payment was made. Also, carriers were required to create four Part B Medicare Annual Data (BMAD) files each year. These three contractor submissions formed the basis of the claims related data available at HCFA.

There were a number of difficulties associated with this data flow. The creation of multiple data files was costly and time consuming. There was a considerable time lag between the rendering of services and the availability of information at HCFA. This time lag was most prominent in the BMAD data files which were only submitted annually, and approximately six months after the end of the calendar year reported. Moreover, since contractors' first priority is tasks related to payment of claims, submission of information to HCFA at times suffered due to operational crises at the contractors.

Payment records contained only summary data. That is, individual procedures on a claim were "rolled up" into a single line item. Thus, the procedure code and the finan-

cial data did not directly correspond, making procedure specific analysis from payment records impossible. In addition, payment records were not submitted on claims in which no payment was made either due to application of the deductible or denial of the claim.

UNIBILL records were not as problematic as Part B Data. Nonetheless, there were problems in records that failed HCFA edits and were returned to the intermediary (RTIs) for review and correction. Millions of the RTIs were not corrected and were posted to the database in error after being held for many months without a response from the intermediary, thus, contributing to the timeliness and quality problems in the old data environment.

Finally, there was little linkage between the data sources. Although the Medicare Automated Data Retrieval System (MADRS) integrated Part A and Part B summary claims information, there was not a direct link to detailed Part B information or the other data sources, such as provider files.

COMMON WORKING FILE (CWF)

We have restructured our claims data systems to provide a more efficient mechanism for claims processing, known as the Common Working File (CWF). CWF is a decentralized Medicare claims validation and benefit authorization process. Under CWF, each Medicare beneficiary is assigned to one of nine host sites. Each host then becomes the repository for all of the official information pertaining to that beneficiary. That is, all beneficiary entitlement information, such as Part A and Part B entitlement dates, Medicare secondary payer (MSP) information, group health plan election and other third party payer information is housed at the host site.

Once a Medicare claim is filed for the beneficiary, the contractor forwards the claims to the host for authorization. The host performs a number of entitlement and uti-

lization edits on the claims received from the contractor. For example, the host performs Part A/B cross over edits, MSP checks, and duplicate claim checks. If the claim is appropriate for payment, authorization is returned to the contractor. If the claim does not pass an edit, the host modifies reimbursement and/or coverage as appropriate and authorizes payment, or notifies the contractor of necessary action. Payment may not be made until the action is taken.

At the time the host authorizes payment, the beneficiary's utilization history is updated. CWF also transmits this information to BDMS. Currently, BDMS also acts as a conduit in processing out of service area claims. That is, when a CWF host receives a claim for a beneficiary not assigned to that host, it transmits the information to BDMS. BDMS, in turn, either forwards entitlement information to the referring host (if it was the first claim ever filed by the beneficiary) or identifies and transmits the claim to the appropriate host responsible for the beneficiary's records.

In summary, CWF is a significant enhancement to Medicare claims processing and ensures uniformity among our contractors in processing claims. By replicating contractor and central office edits, it increases program safeguards and reduces overpayments. Since the CWF software is maintained by a national vendor, program changes can be implemented uniformly and quickly. Finally, information is available to BDMS much more quickly.

NATIONAL CLAIMS HISTORY DATABASE (NCHD)

CWF provides an unprecedented opportunity for the development of a National Claims History Database. Such a database can be used by HCFA to: (1) support improved policy development, both coverage determinations and rate setting; (2) provide enhanced support to program

research and demonstrations; (3) provide essential support to the measurement of quality of care through the measurement of geographic variation and through beneficiary and provider profiling, and (4) support improved program monitoring of contractor activities.

The CWF claims data will be stored in nearline and M204 databases collectively referred to as the National Claims History Database. NCHD encompasses receipt of the claims data, quality control and edits of the data, and creation and maintenance of on-line, nearline, and off-line claims databases, summary databases, a program liability database, and a control file.

NCHD will significantly enhance the information available to HCFA for program monitoring and policy development. First, unlike the previous system which involved considerable delays in obtaining utilization information, NCHD will be available upon payment authorization. In most cases, information will be sent to BDMS prior to payment being mailed from the contractor site. We anticipate having information available to users within a month of claims processing.

In addition, NCHD will house 100 percent of the claims processing information. Not only will NCHD include all of the Medicare claims processed, including those for which denial of payment resulted, but it will also include all of the information provided on the claims, including line item information on all services. Previous data gathering methods did not produce complete data for several reasons, including rejects of intermediary claims that appeared questionable, rolled up Part B services, Part B services for which no payment was made, and aggregation of procedure information.

Moreover, BDMS has devised a system for monitoring the quality of the claims data transmitted to the NCHD. The monitoring process will validate data transmission,

duplicate host edits including consistency, duplicate claims and entitlement checks, and compare summary results to expected values derived through trend analysis.

Approximately 600 million claims per year will be processed and stored in the NCHD. It is not cost effective for the HCFA Data Center to house all this information on-line simultaneously. Consequently, in designing the NCHD, BDMS has planned for on-line, nearline and off-line services. This configuration should provide necessary access to users within acceptable response times. All claims files will be controlled and subject to the quality assurance mechanisms. They will then be routed to their appropriate databases.

The NCHD will maintain considerable on-line capacity. Complete claims information for the "five percent plus" beneficiary sample will be housed on-line. This file includes all information on those beneficiaries whose current account number ends in 05, 20, 45, 70, or 95 or who are entitled to Medicare based on End Stage Renal Disease. If beneficiaries were ever part of the sample, they will remain in the file even if the account number or basis of entitlement changes. This on-line file will include data for the current year to date plus the immediately preceding three years.

The largest proportion of Medicare program expenditures is related to inpatient hospital services. Therefore, the on-line portion of the NCHD will also include 100 percent of inpatient hospital and skilled nursing facility data. Like the beneficiary sample file, this inpatient file will include the data for the current year to date plus the immediately preceding three years.

Also included on-line will be summary data. Although this database design is not yet finalized, it is envisioned that the file will include summary information on data elements of general interest to users. For example, we

envision inclusion of aggregated information, such as claims volume and program payments for each claim type, e.g., skilled nursing facility, hospice, outpatient hospital, etc.

The on-line file will include special study sets. These are special files prepared at the request of other parties for research purposes. For example, the Agency for Health Care Policy and Research will utilize special studies files for the effectiveness initiative.

Finally, on-line access will include a query facility to the near-line and off-line files. This query facility will allow users to build abbreviated data sets for analysis of specific aspects of program policy or procedures. For example, a query may allow users to create a sample of claims for a particular diagnosis, for a specific procedure, for beneficiaries within a particular age range, etc. The query facility will also allow users to link files, such as the enrollment and claims history databases, in order to build files based on beneficiary demographics and services.

We envision that the design of the on-line database will be sufficient to meet most needs of routine users. However, the current year claims information plus the immediately preceding three years' data will be housed in near-line databases to fulfill additional information needs. Although not immediately available to users, the near-line information can be accessed at least monthly. The near-line files will contain item information for 100 percent of the claims, including "no-payment" claims.

Also located in the near-line databases will be the beneficiary program liability files. These are records apart from routine CWF claim submittals that are submitted by each CWF host monthly. The program liability record contains information pertaining to the services for which the

beneficiary incurred a liability through the deductible, institutional coinsurance, Part B limitation (such as psychiatric services), etc.

Once data have aged greater than four years, they will be archived to the off-line files. This information will not be readily available to users, but can be produced on an as-needed basis.

As you can see, this repository for the entire universe of detailed claims data, organized into various databases, will facilitate research, policy development, and decision support. Efforts are currently under way to develop highly functional accesses to NCHD and linkages to other databases and resource files needed to turn databases into knowledge-bases.

SUMMARY OF CURRENT EFFORTS

The inpatient reimbursement focus of the past has now largely shifted to Part B services, with the major exception of HCFA's PPS capital costs initiative. Simultaneously, a host of activities intended to expand data availability, improve its quality and usefulness, and focus research on effectiveness of care are in progress.

Chief among these efforts is Physician Payment Reform (PPR). Enormous efforts in the areas of data collection, data processing and data analysis, either by HCFA staff or under HCFA's guidance have gone into this project. PPR is expected to produce a redistribution of Medicare physician reimbursement via the mechanisms of volume performance standards and resource-based relative value pricing. HCFA will monitor the impacts of PPR through the National Claims History Database.

AHCPR patient outcome research teams (PORTS) are now engaged in studying high frequency conditions and procedures using the NCHD to supplement clinical findings.

HCFA is now putting into place a multi-tiered quality assurance system to support the NCHD. Using a combination of transaction flow monitoring and long term trend analysis, HCFA will assure the integrity and reliability of the increasingly crucial Medicare databases.

To improve its knowledge of physician practices and permit tracking of the impacts of Physician Payment Reform, HCFA will use Common Working File and the NCHD to enforce and monitor unique physician identification number (UPIN) reporting compliance. UPINs will allow HCFA to develop comprehensive profiles of individual physicians and specialty practice patterns. HCFA will also take steps to standardize the definitions of services, as with global surgical packages, as well as the carrier procedures used to evaluate and reimburse claims.

One question that is asked frequently is, "Can nurses be identified uniquely?" At the present time, services provided by nurses to Medicare patients are most often paid as part of the institution's or the physician's bill, although several categories of services provided by nurses can be identified in the NCHD and in Part B Medicare Annual Data (BMAD). The services billed to Medicare directly by Certified Registered Nurse Anesthetists (CRNA) and by Nurse/Midwives (NM) can be identified by specific specialty codes (Code 43 for CRNAs; Code 42 for NMs). Effective April 1, 1990, Nurse Practitioners may bill Medicare separately for services they perform in a skilled nursing facility in collaboration with a physician. Those services can be identified by the presence of a special procedure code modifier (QN). In the future, all nurses who can bill Medicare directly for their services may be identified by a UPIN. This will allow us to link the provider to the beneficiary.

PEER REVIEW ORGANIZATIONS (PROs)

As you are aware, PROs are charged with ensuring that health care services delivered to Medicare beneficiaries are necessary, appropriate, and meet standards of quality. Over the next decade, we plan to move the PRO program away from the current manual, judgmental and variable review of hospital records, toward using large databases of clinical information to identify patterns of inappropriate utilization and poor outcomes. We plan to move the program from penalizing providers for single instances of poor care, toward modifying inappropriate behavior and improving practice by sharing information on patterns of care and outcomes with the local medical community. This transformation incorporates the recent Institute of Medicine report, "A Strategy for Quality Assurance in Medicare."

UNIFORM CLINICAL DATA SET (UCDS)

HCFA is now developing information systems that will allow the PROs to examine the quality of care in a much more scientific and comprehensive way. One of HCFA's important tools for transforming the PRO program will be the Uniform Clinical Data Set (UCDS), a data collection and case finding system for inpatient hospital care. UCDS will permit standardization of the initial PRO review process and will allow us to gather, develop, and analyze extensive clinical data. It will eliminate variations in PRO reviews that now occur because of nurse reviewer's subjective judgment. The UCDS will enable PROs, HCFA and researchers to analyze the effectiveness of medical interventions. We will use the data set to find out what really works well, what has marginal value, and what simply does not work.

In the next few years of the transition to a truly outcome-based program, PROs will abstract up to 1600 relevant data elements into the UCDS from each record under re-

view. The UCDS will then subject the abstracted clinical data to a series of decision rules that serve as computerized quality screens, in order to identify cases needing further review by a PRO physician.

In the longer term, the clinical data abstracted by the PRO will allow the PROs to evaluate patterns of care and patterns of outcomes, adjusted for the condition of patients. We plan to equip PROs with the tools to perform this analysis. These data will also provide a rich epidemiologic database, which we expect to be fertile ground for research into outcomes, effectiveness and quality of care. We will be able to link the UCDS database to currently available Medicare data, which could reveal the total care, both pre- and post-hospitalization, provided to a patient whose hospital record is included in the UCDS file.

The computerized quality screens were initially constructed and reviewed by expert panels, and have been pilot-tested by the PROs. We expect the screens to be constantly reviewed, refined and updated with advances in care and in knowledge of which practices result in the best outcomes. AHCPR's research will certainly be a major source of information for these updates. In addition, we expect AHCPR-developed practice guidelines to be adapted to serve as UCDS quality screens.

The first phase of implementation of UCDS began in January 1991. Seven PROs are now using the UCDS. We anticipate that all the PROs will be using the UCDS by late 1993.

CURRENT BENEFICIARY SURVEY

We are currently working on a new project called the Current Beneficiary Survey (CBS). Beginning in the fall of 1991, the Office of the Actuary and its contractor will survey about 12,000 Medicare beneficiaries, interviewing them about the health care they receive, how much that care costs, and who pays for it. The data from this

survey, linked to the National Claims History Database and the Uniform Clinical Data Set, will be used to: (1) monitor effects of changes in the Medicare program, (2) develop reliable information on services not covered by Medicare, (3) develop reliable information on the costs not assumed by Medicare, (4) understand more about the health status and functioning of the people Medicare serves, and (5) study transitions in beneficiaries' lives and how well the Medicare program responds to those transitions.

POSSIBLE FUTURE DIRECTIONS

Clearly future progress, especially if it is to be transferable to the larger private sector, depends on the availability of comprehensive, reliable data. This implies that HCFA will be able to produce nationally uniform data to support the accepted analysis of clinical, health and financial outcomes, and the relation between alternative treatments and those outcomes. Further, it implies the ongoing opportunity for patients, providers, and payers to have access to analyses which can empower informed choice.

For the NCHD to successfully support policy development, it must provide comprehensive and uniform national data, so that comparability within and across various sectors is possible. That is, data uniformity is the common denominator which makes macro-analysis practicable and credible. To do this, data integration must occur both horizontally and vertically with the health care system it reflects. Horizontal integration calls for the uniformity of coding of the patient's experience across the provider and payer. Vertical integration calls for a uniformity of coding of the patient's experience from provider through the payer's system. HCFA and the DHHS are pressing forward on data integration and uniformity through a large number of mechanisms.

HCFA continues to seek out ways to improve its data collection and the tool set it uses to analyze that data for effective decision support. If the UCDS is successful, an analogue for patients treated in outpatient settings offices could be developed. More work is required on the research construct of "episode of care" so that it can be turned into an operational and reimbursement concept. Followups will be required to determine the penetration of AHCPR guidelines into actual practice pattern changes. To adequately support that, provider profiling from the NCHD will be required. HCFA will need to develop geographical information systems (GIS) and knowledge-based (expert) systems to better organize its data and to apply what we learn to daily Medicare operations.

To be fully successful, we also need to relook at some of the barriers to obtaining crucial data and then optimizing its use. Much confusion is created by the varying enumeration systems now in use. Moving to the personal Social Security Number as the true universal identifier for both patients and providers would be a major step towards data integration. As costs continue to fall for microprocessors, we will need to reevaluate the feasibility of using embedded chips on Medicare beneficiary identification cards to assure the portability/availability of key medical records and insurer data. HCFA must also reassess Medicare's program structure to determine if the current use of the inpatient spell of illness concept as well as deductibles and coinsurance constitute significant barriers for our population to access needed care.

AUTOMATING MEDICAL RECORDS

Legislation has been introduced in the House which would require that hospitals adopt automated patient record systems by the year 2000. The Institute of Medicine and the General Accounting Office have both recently issued reports calling for a centralized effort to develop, promote, and facilitate automated hospital patient record systems.

A number of hospitals, HMO's, and health care systems have installed systems that store and manipulate elements of patient clinical data.

We believe that a great deal more data must be available in order for us to do our job, especially in the quality assurance arena. Therefore, we are examining HCFA's role in a health care system that could provide automated clinical data from the provider level.

CONCLUSION

As we move forward with the program, we will continue to build on past successes, while attempting to use information and incentives to mold the shape of the future. The future is dependent on collecting, organizing, and analyzing appropriate data, and then disseminating those data to policy makers, the health care community and the public. HCFA has made great strides already down this path, and is now gearing up for even greater efforts.

Our challenge is to empower the users so that they can work directly with databases to facilitate and speed their research and analysis. Through our combined efforts, epidemiologists and effectiveness researchers will have access to data of an unprecedented breadth, depth, variety and clinical detail. Through data, they will be able to conduct sophisticated investigations into the impact of health care as it is actually practiced.

Better information means improved outcomes and enhanced quality of care. The importance of valid, complete and accurate data cannot be stressed enough in the pursuit of our goal of identifying and promoting health care interventions that are of consistent high-quality, appropriate, and cost effective. Having better information about the relative outcomes of various treatment options allows providers to make better clinical decisions, gives patients the opportunity to become more involved in these decisions, gives payers better information on what they

are paying for, helps health-services managers make better decisions about resource allocation and the acquisition of new technology, and increases competition based on evidence of quality.