February 2004

A Brief Overview of Medicaid Data Sources

The tabulations of Medicaid spending presented in the tables that accompany this document reflect not only different federal fiscal years (2000 and 2002), but also key differences inherent in the data sources behind these tables. There are two main federal data sources for state-level Medicaid spending by type of service: Form CMS-64 and the Medicaid Statistical Information System (MSIS). While the Centers for Medicare and Medicaid Services (CMS) maintain both sources, there are important differences in the way these data are collected and reported. This document provides some explanation of these differences, as well as more general pros and cons of these two sources.

Form CMS-64

Form CMS-64 is a quarterly statement of expenditures for which states are entitled to federal reimbursement under the authority of Title XIX of the Social Security Act (Medicaid). It is a product of the financial budget and grant system within CMS; it tracks expenditures for the quarter being reported and previous fiscal years, as well as collections or refunds received. It is also the vehicle for adjustments made to correct overpayment and underpayment.

Spending reported on Form CMS-64 are tabulations of actual, documented Medicaid expenditures, drawn from source documents such as invoices, cost reports and eligibility records. If a state is unable to document a claim for expenditures made in the current quarter, the claim must be withheld until it can be supported. The state then reports the amount on a future Form CMS-64 as a prior period adjustment. Spending therefore reflects all expenditures made during the quarter, not all services used.

Data from Form CMS-64 have been provided to us through a combination of data files and spreadsheets compiled by CMS staff. They are generally comparable in content to the "Financial Management Reports" available on the CMS Web site.[1] All told, we have data covering nearly 20 years, from federal fiscal year 1984 to 2002.[2] Expenditures are reported for each fiscal quarter, and are further broken out into more than 40 different types of services and special payment lines. The files also contain information on administrative costs and accounting adjustments.

Data Strengths

One advantage of Form CMS-64 is that these data are more current than data from MSIS. Based on discussions with many state officials, it is clear that states have long-standing procedures to compile and submit these data. Many of these procedures are automated within the states' own systems for tracking and reporting spending, and CMS indicates that all states now submit their CMS-64 data electronically. Further, states' rational desire to receive federal matching funds gives officials incentive to document and report expenditures in a timely manner. All of these factors contribute to the relative alacrity with which these data are collected and, in turn, are made available by CMS. Form CMS-64 is the most timely public source that allows one to observe how

[1] See http://cms.hhs.gov/medicaid/mbes/ofs-64.asp. Last accessed November 10, 2003.
[2] The federal fiscal year begins October 1 and runs until September 30 of the following year.

much money both states and the federal government spent on Medicaid in a given fiscal year, and, perhaps more importantly, to observe trends by type of service.[3]

Another strength of Form CMS-64 is its relative completeness, at least in terms of reflecting virtually all of a state's Medicaid spending. Form CMS-64 is the basis for determining states' matching funds, so it includes all expenditures made under the authority of Title XIX. These include payments to providers for medical services provided to Medicaid enrollees, disproportionate share hospital (DSH) payments,[4] supplemental payments,[5] administrative costs, and adjustments for prior periods. These data are subject to audit and, as such, should be documented. Many researchers view these data as the most reliable source of Medicaid expenditures that can be used to observe trends at the national, state, or service level.

Data Weaknesses

The main limitation of Form CMS-64 is its limited scope. It is not intended for research purposes, but rather is a vehicle for tracking aggregate payments by type of service and determining federal reimbursements. As such, it collects no information concerning the utilization driving these expenditures, such as the number of inpatient hospital visits, prescriptions filled, or nursing home days of care provided. It also does not collect information on the number of individuals using these services or their characteristics (e.g., age, race, gender, or reason for enrollment in Medicaid).

Another limitation of Form CMS-64 is the inability to determine what services are paid for by health plans that receive capitated payments to provide services to Medicaid enrollees, and how much these plans spend for each particular type of service covered. Form CMS-64 includes lines that identify payments to managed care and other health plans, but that is the extent of information available. Expenditures by service on the rest of the form only include payments made under fee-for-service arrangements.

The Medicaid Statistical Information System (MSIS)

MSIS is the principal source of state-reported data on Medicaid enrollees and expenditures. A replacement for Form HCFA-2082, MSIS is a far more comprehensive data source for Medicaid than was previously maintained at a national level.[6] For each person enrolled in Medicaid, MSIS collects information pertaining to why the person is

[3] It is possible to track federal outlays for Medicaid through treasury reports that are released to the public a couple of weeks after the end of each month. However, all that these reports show are total federal outlays during the month and year to date.

[4] Disproportionate share hospitals receive additional Medicaid reimbursement because they serve a higher share of Medicaid (and low-income) patients. States determine if hospitals meet the criteria to be considered a "disproportionate share hospital" and determine the amount of the payment, subject to certain minimum federal standards.

[5] Many states make extra payments to targeted providers, generally nursing homes or hospitals. These enhanced payments can be well in excess of the actual cost of medical services provided to Medicaid beneficiaries as long as the state does not exceed the Medicare upper payment limit, or UPL, which stipulates that Medicaid payments can be no higher than the amount that Medicare would pay for the same service. Whether payments exceed the UPL is based on the aggregate amount that can be paid to an entire class of providers, assuming that every provider in that class was paid Medicare rates for all services provided to Medicaid beneficiaries.

[6] Form HCFA-2082 was an annual report submitted by states to CMS (formerly the Health Care Financing Administration, or HCFA). This report contained highly aggregated data on Medicaid enrollees, use of services, and expenditures for those services. States summarized and reported the data using their own data systems unless they opted to participate in the then-optional MSIS, in which case the 2082 report was produced by CMS using submitted data tapes.

eligible for Medicaid (e.g., do they qualify under rules that apply because of age (65+) or disability, is their eligibility tied to receipt of cash assistance or other standards), which months during the year he/she was enrolled, whether he/she is also enrolled in Medicare or other health insurance, and basic personal characteristics (e.g., age, race, gender). MSIS also collects claims data that are used to generate measures of utilization and payments for each individual.

Spending amounts reported in our MSIS source files reflect payments made for claims adjudicated during the fiscal year, including capitated payments for managed care plans. Like Form CMS-64, this means that payments shown in MSIS are not always tied to services used during that quarter or fiscal year. Unlike Form CMS-64, states generally do include payments made outside the claims processing system (e.g., DSH payments, supplemental payments, and administrative costs).

States have been required to submit data for MSIS to CMS on a quarterly basis since January 1999.[7] CMS and its contractors screen these data using benchmarks intended to achieve a practical level of quality and consistency. Information from approved tapes is entered into MSIS, while rejected tapes must be resubmitted.

While CMS eventually plans to make MSIS data available to a wider audience, public access is currently limited to a few tables for each state that are posted on the CMS Web site.[8] Through an agreement with CMS, we obtained access to data sets drawn from MSIS data that were initially created for use by CMS and congressional agencies. These data include limited information on personal characteristics, eligibility information, service utilization, and medical assistance payments for every person participating in the Medicaid program in FFY 2000.[9] Spending is grouped into 29 different categories.

Data Strengths

The biggest strength of MSIS is its comprehensiveness and, consequently, its adaptability to a wide range of analyses. Even with our source data sets, which are far less detailed than the source MSIS data, it is possible to divide enrollees into many different groupings (e.g., by age, race, or gender, or a combination of these), and then to observe whether they use a particular type of medical service and the aggregate level of payments for those services. With the full range of MSIS data, even more detailed payment analyses are possible. For example, one could observe utilization and payments for specific classes of drugs, or determine spending per day in a hospital or nursing home for people in different groups.[10]

Data Weaknesses

One weakness of MSIS is that its size and intricacy make it susceptible to delays in submission by states and availability from CMS. Historically, statistical information for Medicaid (namely Form HCFA-2082) takes longer to be compiled than expenditure data from Form CMS-64. Delays lengthened with the switch to MSIS, although longer delays were expected simply due to the considerable logistical challenges to be overcome at both the federal and state levels in implementing such a large data system. MSIS was new to many states in FFY 1999, and even those states that participated in earlier

[7] This requirement was enacted as part of the Balanced Budget Act of 1997.
[8] See http://cms.hhs.gov/medicaid/msis/msis99sr.asp. Last accessed November 21, 2003.
[9] Our source data also include some payments that are not attributed to specific individuals and others that are attributed to individuals who are not eligible for Medicaid during this period, but presumably were eligible at one point in time.
[10] The current tape specification and data dictionary for MSIS can be found on the CMS Web site at http://cms.hhs.gov/medicaid/msis/default.asp. Last accessed November 24, 2003.

versions of the system had to adjust to the new process. Hopefully the process will speed up as CMS continues to develop MSIS infrastructure and methods, and the states become more familiar with the system.

A second weakness is also related to the comprehensiveness of MSIS data: with so much information to collect and synthesize, there is a greater likelihood that data will be missing or erroneous. CMS and its contractors also run tests to validate data, check for consistency of data elements over different periods, and verify reasonableness of data. However, the general quality of the data is only as reliable as the data submitted to CMS by states. Some states have been unable to provide some data during the first few years under the new system, including FFY 2000. Some payments are known to be missing or inaccurate in our source files and additional errors may be introduced when CMS compiles data into more aggregated data sets or standardized reports for each state (e.g., through an unnoticed mistake in computer programming). As a result, the tabulations shown in the accompanying tables come with an extensive list of caveats, which are posted as a separate document.

One last weakness of MSIS that deserves note is that it does not include all types of Medicaid expenditures. This "weakness" is perhaps more appropriately described as a "substantive difference" between MSIS and Form CMS-64. Total MSIS payments do not agree with the CMS-64 financial figures because they do not include payments made outside the claims processing system, such as DSH payments. Payments to DSH hospitals typically do not appear in MSIS since states directly reimburse these hospitals and there is no fee-for-service billing. Likewise, the supplemental (UPL) payments discussed above are less likely to appear in MSIS. Other differences may occur because states compile and report these data through separate systems, and there may be differences in accounting practices within these systems.

Discussion

The regrettable but honest conclusion that we draw from many years' experience working with data from Form CMS-64 and our recent forays into MSIS is that neither of these data sources is ideal. This documentation and the notes that go along with the accompanying tables lay out some of the most important differences, but they are by no means exhaustive of all the issues involved in using these data. However, both data sources provide valuable insights into the Medicaid program and we hope users will find the accompanying tables useful and informative.

By way of conclusion, some generalizations are worth mentioning. First, not all of the spending claimed on Form CMS-64 applies to Medicaid enrollees, particularly the amounts reported for inpatient hospital services and nursing facilities, which are inflated in some states by UPL payments. Second, the less aggregated the results, the more prone they are to error, particularly when MSIS is the source. In other words, national results are stronger than state-level results, and more inclusive tabulations (e.g., total spending) are stronger than results for individual groups (e.g., children) or services (e.g., inpatient hospital). Lastly, federal Medicaid data sources are best used to observe national trends and to make broad comparisons across states. More detailed, state-level analyses—especially those that intend to focus only on a single state—may be better when conducted using data obtained directly from the state itself.