United States General Accounting Office

GAO

Report to Congressional Committees

May 2002

MEDICARE

Recent CMS Reforms Address Carrier Scrutiny of Physicians' Claims for Payment

GAO-02-693

Contents

Tables

Figure

Abbreviations

CMS	Centers for Medicare and Medicaid Services
CERT	Comprehensive Error Rate Testing
CPE	contractor performance evaluation
D.O.	doctor of osteopathy
E&M	evaluation and management
HCFA	Health Care Financing Administration
HHS	Department of Health and Human Services
LMRP	local medical review policy
M.D.	doctor of medicine
NHIC	National Heritage Insurance Company
OIG	Office of Inspector General
PCA	Progressive Corrective Action
PIMR	Program Integrity Management Reporting system
PIN	provider identification number
PSC	program safeguard contractor
WPS	Wisconsin Physicians Service Insurance Corporation

GAO
Accountability * Integrity * Reliability

United States General Accounting Office
Washington, DC 20548

May 28, 2002

Congressional Committees

In 1990, we designated the Medicare program to be at risk of considerable losses to waste, fraud, and abuse because of its vast size, complex structure, and weaknesses in both financial and program management. More than a decade later, we still consider Medicare to be a high-risk program.[1] With annual fee-for-service payments now totaling about $192 billion, Medicare finances health services delivered to elderly and disabled individuals by hundreds of thousands of providers. The Centers for Medicare and Medicaid Services (CMS)[2]—the federal agency that manages the Medicare program—is responsible for ensuring that these funds are spent appropriately. However, the process of enforcing program payment rules has raised concerns that the impact of these safeguard activities has imposed too great a burden on health care providers.[3]

With an interest in striking a balance between appropriate payment controls and reasonable billing requirements for providers, the Congress required, in the Medicare, Medicaid and SCHIP Benefits Improvement and Protection Act of 2000, that we study Medicare claims review—designed to detect improper billing or payment—and related education activities for physicians.[4] While CMS contractors responsible for processing physicians' Medicare claims—referred to as carriers—conduct an automated check of all claims submitted, they select only a sample of claims for medical review. For the purposes of our study, these are reviews that involve a detailed examination of claims by clinically trained staff and require that

[1]U.S. General Accounting Office, *High Risk Series: An Update*, GAO-01-263 (Washington, D.C.: January 2001).

[2]Until June 14, 2001, CMS was known as the Health Care Financing Administration (HCFA).

[3]In June 2001, we responded to questions raised by the Senate Finance Committee that were related to these concerns. See U.S. General Accounting Office, *Regulatory Issues for Medicare Providers*, GAO-01-802R (Washington, D.C.: June 11, 2001).

[4]Pub. L. No. 106-554, App. F, Sec. 437(a), 114 Stat. 2763A-463, 2763A-527. Although Medicare considers services from dentists, optometrists, podiatrists, and chiropractors to be covered physicians' services (see 42 C.F.R. § 410.20(b)(2002)), as agreed with the committees of jurisdiction we focused on claims filed by doctors of medicine (M.D.s) and doctors of osteopathy (D.O.s) only.

physicians submit medical records to substantiate their claims for payment. In fiscal year 2001, CMS revised its policy on medical reviews of physicians' claims under the Progressive Corrective Action (PCA) initiative, directing carriers to focus their scrutiny on claims where there is the greatest risk of inappropriate payments.[5]

As agreed with the cognizant congressional committees, we focused our study on the medical review process and the related implications of PCA's implementation. Specifically, we examined (1) the extent to which physicians have claims that are subjected to medical review, (2) the implications for physicians of PCA's strategic approach to overpayment assessments and education, (3) the accuracy of carriers' decisions to pay or deny a claim based on medical review, (4) the effectiveness of criteria used to identify claims for medical review that have potential billing errors, and (5) how CMS evaluates carrier efforts to reduce physicians' billing errors.

Our study covers medical review activities, excluding fraud-related cases, conducted largely in fiscal year 2001. Because national data specific to medical reviews of physicians' claims were not available, we contacted three carriers to obtain information that is only maintained at the carrier level. These carriers are National Heritage Insurance Company (NHIC) in California, Wisconsin Physicians Service Insurance Corporation (WPS), and HealthNow NY; they serve six states and process claims for about one-quarter of Medicare's participating physicians.[6] We interviewed carrier officials about their selection of claims for medical review, the medical review process, and related communication with physicians. In addition, we collected data on physician practices that had claims subjected to medical review, overpayment assessments, and requests for repayment extensions. We also interviewed officials at CMS's central and regional offices and representatives of physician associations in several states.

[5]Department of Health and Human Services, HCFA, *Medical Review Progressive Corrective Action*, Program Memorandum Transmittal AB-00-72 (Baltimore, MD: Aug. 7, 2000).

[6]These carriers vary by size and geographic region. NHIC's California component is a large insurer with separate facilities serving the southern and northern areas of the state. In some instances, data for fiscal year 2001 did not include the entire year for NHIC because its southern office did not assume carrier operations until 2 months after the fiscal year had begun. Prior to December 2000, another carrier conducted claims review for southern California. WPS, also a large insurer, has separate facilities that operate in four states (Wisconsin, Illinois, Michigan, and Minnesota). The Minnesota office was the most recent addition, joining WPS in September 2000. HealthNow NY is a small insurer that serves providers in upstate New York.

In addition, we contracted with a firm with expertise in Medicare's medical review activities to independently assess the accuracy of the three carriers' medical review decisions. Its findings were discussed with carrier officials and a consensus was reached on the correct medical review decision in all but one case. The accuracy of the carrier decision in that case was decided by the acting deputy director of CMS's Program Integrity Group—a physician. (For a detailed description of the validation process, see app. I.)

Because the study was limited to three carriers, our findings regarding the frequency and accuracy of claims reviews cannot be generalized to the universe of carriers. The carriers performed a series of special analyses to provide data necessary for our study and experienced varying degrees of difficulty in extracting data from their information systems. Because the data are maintained in multiple systems and in various formats, some information was not readily available and could not be included in the tables we present. We did not verify the accuracy or completeness of the data provided by the carriers. Also, although part of the study's mandate, as agreed with committee staff we did not assess the adequacy of resources that CMS devotes to physician education regarding the claims review process. CMS policy changes concerning the focus of physician education have been too recent to allow for analysis of the sufficiency of related resources. We performed our work from June 2001 through March 2002 in accordance with generally accepted government auditing standards.

Results in Brief

Our review at three carriers indicates that most physicians billing Medicare are largely unaffected by carriers' medical reviews. In fiscal year 2001, at least 90 percent of physician practices had no claims subjected to a medical review. The share of physician practices that had any claims subject to medical review before payment was 10 percent in states served by the Wisconsin carrier and a smaller proportion in California and upstate New York. For the typical practice, the carriers reviewed 2 claims during the year. One-tenth of 1 percent of physician practices had claims selected for medical review after they were paid. These reviews typically involved about 30 to 50 claims.

At our three carriers, implementation of PCA has effectively reduced the amounts that physicians must repay Medicare based on medical reviews of their claims, and has increased carrier education to individual physicians. Under PCA, carriers must limit their use of extrapolation—a process by which carriers estimate the amount Medicare overpaid a practice by

projecting the error rate found in a sample of its claims—to those cases that involve major billing problems. In fiscal year 2001, the carriers in our study virtually eliminated extrapolation. Following this and other modifications related to PCA, the highest overpayment amounts assessed physician practices decreased substantially. In addition, the three carriers increased direct education and feedback to physicians concerning the results of medical reviews and proper billing practices so that future claims would be submitted correctly.

With relatively few exceptions, the carriers in our study made appropriate payment determinations in examining the physician claims selected for a medical review. Our contractor's evaluation of the carriers' medical review decisions found a 96 percent overall accuracy rate. The accuracy of carriers' decisions to totally deny payment was even higher, 98 percent. For reviews where the carrier paid a reduced amount on a physician's claim, the accuracy of carrier decisions was somewhat less—92 percent. Such reductions occurred most often on claims reviews involving what should have been the appropriate billing level for physician office visits. Overall, the small share of inaccurate decisions made by the carrier resulted in both overpayments and underpayments.

While the three carriers were highly accurate in their payment decisions, they could improve their selection of claims for medical review by better identifying claims likely to have been billed incorrectly. Fiscal year 2001 data showed substantial variation in the performance of edits—criteria used to target specific services for review—that our three carriers employed to identify medically unnecessary or incorrectly coded physician services. For the prepayment edits that accounted for the largest number of claims examined by each of our carriers, denial rates—that is, the proportion of reviewed claims that were fully or partially denied—ranged from 5 to 82 percent. By refining their selection criteria to more consistently target claims likely to have been submitted with errors, carriers could improve the efficiency of their own operations and reduce administrative demands on the small proportion of physician practices with claims selected for review.

CMS is refocusing its oversight of carrier performance in processing and reviewing claims. Specifically, the agency intends to hold carriers accountable for the overall level of payment errors in all the claims they process, not just the ones they review. Consistent with this approach, CMS is developing a new tool—the Comprehensive Error Rate Testing (CERT) program—that involves having an independent contractor determine the accuracy of claims processed and paid by each carrier using quantitative

performance measures. CMS expects CERT to help identify individual carrier performance problems and track each contractor's rate of improvement. CERT benchmarking is expected to be in place by November 2002.

We provided CMS a draft of this report for comment. The agency generally agreed with our findings.

Background

CMS, within the Department of Health and Human Services (HHS), provides operational direction and policy guidance for the nationwide administration of the Medicare program. It contracts with private organizations—called carriers and fiscal intermediaries—to process and pay claims from Medicare providers and perform related administrative functions. Twenty-three carriers nationwide make claims payments for physician services, which are covered under part B of Medicare.[7] In addition, carriers are responsible for implementing controls to safeguard program dollars and providing information services to beneficiaries and providers. To ensure appropriate payment, they conduct claims reviews that determine, for example, whether the services physicians have claimed are covered by Medicare, are reasonable and necessary, and have been billed with the proper codes.

Carriers employ a variety of review mechanisms. Automated checks, applied to all claims, are designed to detect missing information, services that do not correspond to a beneficiary's diagnosis, or other obvious errors. They may also be used to determine if a claim meets other specific requirements, including national or local coverage policies (such as allowing only one office visit for an eye examination per beneficiary per year unless medical necessity is documented).[8] Manual reviews by carrier staff are used when the review of a claim cannot be automated to

[7]Part B also covers charges from licensed practitioners, as well as clinical laboratory and diagnostic services, surgical supplies and durable medical equipment, and ambulance services. Part A covers hospital and certain other services.

[8]Local coverage rules, known as local medical review policies (LMRPs), reflect regional differences in medical practice by specifying the circumstances under which a carrier will or will not provide Medicare payment for a particular service and how the service will be coded. According to CMS officials, LMRPs are carriers' interpretations of Medicare coverage for a particular service that enhance or clarify national Medicare policy or provide guidance in the absence of national policy. Because these interpretations may differ, one carrier might pay for a particular service that would not be paid for by another carrier.

determine if sufficient information has been included to support the claim. In the most thorough type of manual claims review, a carrier's clinically trained personnel perform a medical review, which involves an examination of the claim along with the patient's medical record, submitted by the physician, to determine compliance with all billing requirements.

Typically, carriers conduct medical reviews on claims before they are paid, by suspending payment pending further examination of the claim. Prepayment medical reviews help to ensure that a carrier is making appropriate payment decisions while the claims are processed, rather than later trying to collect payments made in error. To target such reviews, carriers develop "edits"—specific criteria used to identify services that the carrier determines to have a high probability of being billed in error. Carriers develop these edits based on data analyses that include comparisons of local and national billing patterns to identify services billed locally at substantially higher rates than the national norm.[9] Carriers may also develop edits for prepayment medical review based on other factors, such as CMS directives or individual physicians or group practices the carrier has flagged for review based on their billing histories. Before putting edits into effect, CMS expects the carriers to conduct targeted medical reviews on a small sample of claims in order to validate that the billing problem identified by the carrier's data analysis or other sources does actually exist.

In addition to prepayment medical reviews, carriers conduct some medical reviews after claims are paid. Postpayment reviews determine if claims were paid in error and the amounts that may need to be returned to the Medicare program. They focus on the claims of individual physicians or group practices that have atypical billing patterns as determined by data analysis. Such analyses may include comparisons of paid claims for particular services to identify physicians who routinely billed at rates higher than their peers. Carriers may also select claims for postpayment review based on other factors, such as information derived from prepayment reviews, referrals from other carrier units, and complaints from beneficiaries. In rare cases, postpayment reviews may result in referrals to carrier fraud units.

[9]CMS maintains reports containing national averages for the billing of specific services. Billing data are also available by physician specialty, locality, and other categories.

Each year, as part of their budget negotiations with CMS, carriers develop medical review strategies that include workload goals for conducting medical reviews. CMS provides each carrier with an overall budget for claims review. The carriers then submit for CMS approval their workload goals for specific activities, such as the number of prepayment and postpayment medical reviews they plan to conduct, along with proposed budgets and staff allocations across these activities. In addition, the carriers submit budget proposals for provider education and training related to issues identified in medical review. CMS requires the carriers to reassess the allocation of these resources among review and educational activities during the course of the year and, with CMS approval, to shift resources as appropriate to deal with changing circumstances.

Few Physician Practices and Few Claims Per Practice Receive Medical Reviews

In estimating the prevalence of medical reviews, data from the three carriers in our study show that more than 90 percent of physician practices—including individual physicians, groups, and clinics—did not have any of their claims selected for medical review in fiscal year 2001, and for those that did, relatively few claims were subject to review.

A small proportion of physician practices served by the three carriers had any claims medically reviewed during fiscal year 2001. Table 1 shows that about 10 percent of the solo and group practices that filed claims with WPS had any prepayment medical reviews. This proportion was even lower at HealthNow NY and NHIC California, with rates of about 4 and 7 percent, respectively. The share of physician practices with postpayment reviews by any of these carriers was much smaller; approximately one tenth of 1 percent of practices had claims selected for medical review after payment had been made.

Table 1: Physician Practices Whose Claims Received Medical Review, Fiscal Year 2001

	NHIC California[a]		WPS[b]		HealthNow NY	
Medical review	Number	Percent of total[c]	Number	Percent of total[d]	Number	Percent of total[d]
Prepayment	5,590	7.4	13,732	10.1	1,270	4.3
Postpayment	113	0.1	80	0.1	33	0.1

Note: Physician practices were identified by the Medicare Provider Identification Number (PIN).

[a]The number of practices shown include data from northern California for November 2000 to September 2001 and from southern California for December 2000 to September 2001.

[b]WPS prepayment data include reviews in Illinois, Michigan, and Minnesota only; data were not available for Wisconsin. Postpayment data include Illinois, Michigan, Minnesota, and Wisconsin.

[c]Because a list of active PINs was not available from NHIC California, we estimated the total number of solo and group practices in California based on data from the most recent American Medical Association census of group medical practices, adjusted for increases in the total number of nonfederal M.D.s as of December 31, 2000, and the number of D.O.s in the state.

[d]Percentages are based on lists of active PINs obtained from the carrier.

Source: GAO analysis of carrier data, and physician practice data from the American Medical Association and American Osteopathic Association.

Further, for most of the physician practices having any claims subject to medical review in fiscal year 2001, the carriers examined relatively few claims. As shown in table 2, over 80 percent of the practices at each carrier whose claims received a prepayment review had 10 or fewer claims examined and about half had only 1 or 2 claims reviewed.

Table 2: Number of Claims Per Physician Practice Subject to Prepayment Medical Review, Fiscal Year 2001

	Percent of practices whose claims were reviewed		
Claims per practice	NHIC California[a]	WPS[b]	HealthNow NY
1 or 2	56.4	54.0	50.9
10 or fewer	86.4	88.7	81.7
100 or more	1.2	0.5	2.9

[a]The figures shown include data from northern California for November 2000 to September 2001 and from southern California for December 2000 to September 2001.

[b]WPS prepayment data include Illinois, Michigan, and Minnesota; prepayment data were not available for Wisconsin.

Source: GAO analysis of carrier data.

For the small number of physician practices whose claims were subject to postpayment review in fiscal year 2001, the three carriers typically examined more claims per practice. At NHIC California, the median

physician practice had 33 claims reviewed postpayment; at WPS, 49; and at HealthNow NY, 31.

New CMS Review Policy Has Reduced Physician Repayment Amounts Due and Increased Focus on Physician Education

With the issuance of the PCA initiative, CMS modified the approach that carriers use to select physicians' claims for medical review, determine repayments due, and prevent future billing errors. PCA directs carriers to (1) use their analyses of physician billing patterns to better focus their medical review efforts towards claims with the greatest risk of inappropriate payments, and (2) provide targeted education regarding how to correct billing errors. Information from our three carriers indicates that, as a result of PCA, they virtually eliminated in fiscal year 2001 their use of extrapolation, a corrective action that involves projecting a potential overpayment from a statistical sample. A recent CMS survey also showed reduced use of extrapolation by other carriers. After PCA was implemented, the highest repayment amounts each of our three carriers assessed physicians were substantially lower than in the previous year. The carriers have also developed medical review strategies that include increased education for individual physicians in an effort to change billing behavior and, thus, prevent incorrect payments.

CMS Policy Matches Corrective Actions to Level of the Physician's Billing Problems

PCA seeks to more effectively select physician claims for medical review. The initiative aims to further the agency's program integrity goals of making sure that claims are paid correctly and billing errors are reduced while carriers maintain a level of medical review consistent with their workload agreements with CMS. In targeting physician claims, PCA requires that carriers subject physicians only to the amount of medical review necessary to address the level and type of billing error identified. If claims data analysis shows a potential billing problem for a particular service, carriers must first conduct a "probe review"—requesting and examining medical records from a physician for a limited sample of claims—to validate suspicions of improper billing or payment. For example, a carrier may initiate a postpayment probe review after discovering that a physician billed, per patient, substantially more services than his or her peers.[10] If the carrier determines that the documentation in

[10]Provider-specific probe reviews can be both prepayment and postpayment. If the carrier can select a sufficient number of claims for a probe in a reasonable period, it may choose to conduct a prepayment medical review. For lower volume services, however, the carrier will typically take a postpayment approach so that the physician does not have an excessive wait before having claims processed.

the medical records does not support the type or level of services that was billed, the carrier calculates an error rate—the dollar amounts paid in error relative to the dollar amount of services reviewed. The error rate, the dollar value of the errors, and the physician's past billing history are among the factors the carrier may consider in assessing the level of the billing errors and determining the appropriate response.[11]

Under PCA, CMS instructs carriers to categorize the severity of billing errors found in probe samples into three levels of concern—minor, moderate, or major. Minor concerns may include cases with a low error rate, small amounts improperly paid, and no physician history of billing problems. Moderate concerns include cases that have a low error rate but substantial amounts improperly paid. Major concerns are cases with a very high error rate, or even a moderate error rate if the carrier had previously provided education to the physician concerning the same type of billing errors. Although no numerical thresholds were established in the instructions to carriers, CMS provided vignettes illustrating the various levels of concern. In an example of a major concern, 50 percent of the claims in a probe sample were denied, representing 50 percent of the dollar amount of the claims reviewed.

PCA allows carriers flexibility in determining the most appropriate corrective action corresponding to the level of concern identified. At a minimum, the carrier will communicate directly with the provider to correct improper billing practices. For probe reviews that are conducted postpayment—the stage at which probe reviews are most commonly done at the three carriers we visited—they must also take steps to recover payment on claims identified as having errors. Further options for corrective action include:

- for minor concerns, conducting further claims analysis at a later date to ensure the problem was corrected;
- for moderate concerns, initiating prepayment medical review for a percentage of the physician's claims until the physician demonstrates compliance with billing procedures; and
- for major concerns, initiating prepayment medical review for a large share of claims or further postpayment review to estimate and recover potential

[11]PCA identifies secondary considerations that carriers should use in determining appropriate corrective actions. Aggravating factors might include past history of abusive billing practices or a high percentage of particular types of errors. Mitigating factors include establishing a compliance training program for office staff.

overpayments by projecting an error rate for the universe of comparable claims—a method of estimation called "extrapolation."

Under PCA, because the corrective action is scaled to the level of errors identified, the potential financial impact of medical review on some physicians has decreased. Although our three carriers did not frequently use extrapolation in 2000, before PCA, a physician could experience a postpayment medical review that involved extrapolation regardless of the level of errors detected. As shown in table 3, after PCA's implementation, the highest amount any physician practice was required to repay substantially declined at the three carriers. The largest overpayment assessed across the carriers ranged from about $6,000 to $79,000 in fiscal year 2001, compared with about $95,000 to $372,000 in the previous year. At the same time, changes in the median overpayment amounts varied across our three carriers, with a dramatic decline at NHIC California. (Recovery of overpayments from physicians is discussed in app. II.)

Table 3: Overpayments Assessed Physician Practices, Fiscal Years 2000-2001

	NHIC California[a]		WPS		HealthNow NY	
	2000	**2001**	**2000**	**2001**	**2000**	**2001**
Number of practices assessed an overpayment	58	81	106[b]	76[b]	158	151
Overpayment per practice[c]						
Median amount	$11,644	$2,023	$2,185	$2,913	$134	$133
Highest amount	174,838	79,313	94,545	79,488	372,446	6,449

Notes: Overpayment assessments can result from billing errors found in one or more claims or be extrapolated from errors found in a sample of claims. Some overpayment assessments may reflect the outcomes of medical reviews conducted the previous fiscal year.

[a]The figures shown include data from NHIC's northern California office only; data were not available for its southern office for fiscal year 2000. However, during fiscal year 2001, the southern California office's median overpayment assessment was $101 and the highest amount was $18,396.

[b]WPS data represent the number of overpayment assessments. Because a few physicians were assessed more than one overpayment during the fiscal year, these data very slightly overstate the number of physician practices.

[c]Some assessments may have been subsequently reduced after an appeal.

Source: NHIC California, WPS, and HealthNow NY.

Several factors may account for the lower overpayment amounts assessed physician practices in fiscal year 2001. Under PCA, probe samples are designed to include a small number of claims per physician, so any overpayments discovered through the probe review process will likely be limited. Whereas the typical postpayment medical review conducted before PCA might involve several hundred claims, a probe review

generally samples 20 to 40 claims selected from an individual physician for the time period and the type of service in question. If the carrier classifies the physician's billing problem as a minor or moderate level of concern, the physician is responsible for returning only the amount paid in error found in the probe sample. In these cases, there would not be an extrapolation as may have occurred in the past.

The circumstances in which carriers determine an overpayment by extrapolating from a statistical sample have narrowed. Before PCA was implemented, carriers were encouraged to extrapolate an overpayment amount whenever a postpayment sample of claims was drawn. However, even then, our three carriers used extrapolation in only 38 instances in fiscal year 2000. Now CMS has directed carriers to reserve the use of extrapolation for those cases where a major level of concern has been identified. In addition, before it can proceed with an extrapolation, the carrier has to draw a new, statistically valid random sample from which to project the assessed overpayment.[12] Furthermore, the amount to be recovered based on an extrapolation is smaller than it typically would have been in years past because instead of using the average overpayment found in the sample, the average is reduced because statistical estimates do not have 100 percent accuracy.[13]

In the event that extrapolation is used, the requirement to start with probe samples may also reduce the physician's financial risk. Because a probe sample is fairly small, carrier officials stated that they may only examine one or two types of services, compared to four to six types of services reviewed previously. This means that if the probe review results lead to an extrapolation based on a larger statistically valid random sample, only claims for the small number of service types will be included in that sample and the results will be projected to a smaller universe of claims.

[12]An exception where extrapolation based on the original probe sample is allowed, is when the physician chooses to accept a proposed consent settlement rather than having to submit medical record documentation for a new, and typically larger, sample of claims.

[13]In a typical extrapolation, the amount of the overpayment is calculated by (1) determining the average overpayment per claim in the sample as a whole or broken down into strata or clusters, (2) multiplying that amount by the number of corresponding claims in the universe, and (3) reducing that amount to that represented by the lower bound of a one-sided 90 percent confidence interval. This third step was introduced in January 2001, when CMS issued new standards for statistical sampling and extrapolation methodologies used by carriers. This change takes into account that statistical estimates may be in error and that the actual amount may fall within a range around an estimate. This policy involves using the bottom of the range as the amount of overpayment to recover.

Consequently, the total amount assessed would tend to be smaller than previously extrapolated amounts.

In the first year of PCA implementation, our three carriers virtually eliminated their use of extrapolation to determine overpayments. For example, NHIC California officials stated that before PCA it was not uncommon to use extrapolation in determining overpayments based on samples involving a relatively large number of claims. But now, such extrapolation is to be used infrequently. If a physician failed to correct inappropriate billing practices following a probe sample and targeted education, the carrier would probably subject some or all of the physician's subsequent Medicare billing for prepayment review before it would consider selecting a larger postpayment sample suitable for extrapolation. As shown in table 4, in fiscal year 2000, NHIC California conducted 31 postpayment reviews that involved extrapolation, with a median overpayment assessment of about $32,000, but had no cases involving extrapolation in fiscal year 2001. Similarly, HealthNow NY had none in fiscal year 2001 and WPS reported no cases of extrapolation other than a small number of consent settlement cases.[14]

[14]Under a Medicare consent settlement, a potential overpayment is determined by extrapolating from a small sample of claims that is not statistically valid. The carrier would then offer the provider the option of repaying the projected overpayment and agreeing to a consent settlement or proceeding to a further review of a larger, statistically valid random sample of claims and overpayment projection. Of the carriers in our study, only WPS' Minnesota office used consent settlements for a few cases in either fiscal year 2000 or 2001, and it settled all but one of its cases.

Table 4: Carriers' Use of Extrapolation in Assessing Overpayments to Physician Practices, Fiscal Years 2000-2001

	NHIC California		WPS		HealthNow NY	
	2000	2001	2000	2001	2000	2001
Number of overpayment cases involving extrapolation	31	0	6	0	1	0
Size of claims samples used						
Smallest	43	[a]	60	[a]	[a]	[a]
Median	207	[a]	171	[a]	43	[a]
Largest	1,232	[a]	432	[a]	[a]	[a]
Projected overpayment						
Lowest amount	$3,758	[a]	$2,640	[a]	[a]	[a]
Median amount	32,140	[a]	29,093	[a]	$112,896	[a]
Highest amount	234,890	[a]	72,679	[a]	[a]	[a]

Note: Because a physician practice may have more than one sample of claims selected in a year, overpayments were reported for each case where extrapolation was used. Some projected overpayments were later reduced as the result of physician rebuttals or appeals.

[a]Not applicable.

Sources: NHIC California, WPS, and HealthNow NY.

A recent CMS survey indicates that most carriers limit their use of extrapolation. In October 2001, CMS surveyed carriers to determine, in part, the number of cases that involved extrapolation during the last 3 fiscal years.[15] Of the 18 carriers that responded to the survey, only 3—serving Ohio, West Virginia, Massachusetts, and Florida—had more than 9 cases involving extrapolation in fiscal year 2001.[16]

[15]These cases involved statistically valid random samples of claims that were used to project overpayments. The survey also identified other cases involving consent settlements based on extrapolations from more limited samples of claims.

[16]One of the three carriers did not separate the number of extrapolation cases for medical review from those associated with fraud-unit activity. An official at that carrier told us that approximately 33 of the 131 cases where extrapolation was used were related to medical review.

Carriers Are Expected To Integrate Medical Review And Education Outreach Functions

A key focus of PCA is its emphasis on carrier feedback to physicians in the medical review process. Educating physicians and their staffs about billing rules is intended to increase correct billing, which reduces both inaccurate payments and the number of questionable claims for which physicians may be required to forward copies of patient medical records. When a carrier identifies a physician's billing problem, PCA requires the carrier to provide data to the physician about how his or her billing pattern varies from other physicians in the same specialty or locality. For issues that affect a large number of providers, CMS recommends that carriers work with specialty and state medical societies to provide education and training on proper billing procedures.

In response to PCA, two of the three carriers planned substantial increases in their spending for education and feedback to physicians on medical review issues as part of their overall medical review strategies for fiscal year 2002. As shown in table 5, the three carriers had budget increases of various sizes for provider education and training related to medical review.[17]

Table 5: Carrier Budgets for Provider Education and Training Related to Medical Review Activities, Fiscal Years 2001-2002

	NHIC California[a]	WPS	HealthNow NY
Fiscal year 2001	$491,817	$645,561	$277,939
Fiscal year 2002	767,032	736,000	284,000
Percent change	+56.0	+14.0	+2.2

Note: Data for fiscal year 2001 represent actual expenditures; data for fiscal year 2002 are estimates.

[a]Because NHIC California's southern office did not assume carrier operations until December 2001, fiscal year 2001 includes only 10 months for that office and all 12 months for the northern office. As a result, the percentage change in the budget for fiscal year 2002 is overstated as the budget for that year covers 12 months for both offices.

Source: CMS and NHIC New England.

[17]All education and training activities related to medical review are funded through the Medicare Integrity Program, which also supports claims reviews and antifraud activities. General provider education related to enrollment and billing procedures is funded from a larger and separate budget for program management. The related fiscal year 2002 budgets for the three carriers were: $1.6 million for NHIC California, $2.8 million for WPS, and $1.3 million for HealthNow NY. These represented increases from the previous fiscal year of 42, 6, and 102 percent, respectively.

As part of their strategies to increase physician education, the three carriers reported that they were making greater use of phone calls and individualized letters to physicians' offices to notify them about billing errors. Carriers record their contacts using physician tracking systems to check on the education that has been provided to the physician, which can include letters, materials, phone calls, or face-to-face visits. Whereas in the past it was common for carriers to simply point physicians toward the applicable Medicare rules, under PCA they have assisted physicians in interpreting the rules and applying them to specific billing situations. The carrier's medical review staff has addressed problems of questionable billing patterns by contacting physicians by phone to provide specific information pertaining to billing rules. For physicians whose claims are undergoing postpayment review, the carrier sends a letter at the completion of the medical review that provides a description of the billing problems found, including, as needed, information on the relevant national and local medical policies. The letter also identifies a contact person at the carrier, should the physician want additional information about billing or documentation issues.

For example, WPS officials acknowledged that they previously had little or no follow-up with physician practices whose claims were denied or reduced after medical review to make sure they understood how to bill correctly. In fiscal year 2001, WPS began providing additional education—some efforts addressing all Medicare physicians and some targeted to providers in specific specialties or service locations. To identify the groups that would most benefit from targeted education, the carrier developed benchmark data on billing errors using aggregate claims data on utilization, denial rates, and other billing patterns. For example, the carrier developed education campaigns targeted to mental health practitioners, such as psychologists, clinical social workers, and psychiatrists. In fiscal year 2001, WPS also began to conduct on-site education and group meetings and contact specialty associations to disseminate further information.

Independent Review Confirms Accuracy of Carriers' Payment Decisions

In addition to concerns about having their claims selected for medical review, some physicians have expressed dissatisfaction with the accuracy of the carrier medical review decisions concerning the medical necessity, coding, and documentation of physician services billed to Medicare. To assess the appropriateness of clinical judgments made by carriers' medical review staff, we sponsored an independent evaluation by the private firm that monitors claims payment error rates as a Medicare program safeguard contractor.[18] The firm found that our three carriers made highly accurate medical review decisions. In addition, the level of accuracy was highly consistent across the three carriers. Slight variation in the degree of accuracy was evident when the claims reviewed were classified by the type of payment decision: to pay the claim in full, to pay a reduced amount, or to fully deny payment.

The independent review was conducted on samples of 100 physician claims from each carrier selected randomly from all claims undergoing either prepayment or postpayment medical review in March 2001. Nurse reviewers examined the carrier's initial review decision to see if it was supported by the available medical record documentation and carrier policies in effect when the carrier made its payment decision. These reviewers then discussed with the carrier's staff each claim where they had come to a different conclusion, and in all but one instance, the carrier and contractor achieved a consensus as to whether the original carrier decision was in error. The acting deputy director of CMS's Program Integrity Group, a physician, decided the accuracy of the one case that remained in dispute.

For the vast majority of claims, the independent reviews validated the carriers' decisions. As shown in table 6, the independent reviewers agreed with carriers' original assessments in 280 of the 293 cases examined, or about 96 percent of the time.[19] The small share of inaccurate decisions

[18]HCFA chose 12 claims administration contractors in 1999 to act as program safeguard contractors (PSCs) for Medicare and since then has issued task orders that include different ways of using PSC services. Some task orders involve discrete activities by a single PSC that focus on specific areas vulnerable to fraud and abuse, others require PSCs to replace some or all of the program safeguard activities traditionally performed by claims administration contractors, and still others may have a national impact on fraud and abuse prevention and detection.

[19]This result was consistent across the three carriers and for both prepayment and postpayment reviews. (See app. I.)

made by the carrier resulted in both overpayments and underpayments to physicians.

Table 6: Accuracy of Carrier Medical Review Decisions on Physician Claims

		Inaccurate decision rate	
Carrier decision	Accurate decision rate (percent)	Overpayment (percent)	Underpayment (percent)
All decisions on sampled claims[a] (n=293)	95.6	2.7	1.7
Deny in full (n=64)	98.4	0.0	1.6
Deny in part (n=59)	91.5	1.7	6.8
Pay in full (n=170)	95.9	4.1	0.0

[a]Claims randomly selected from all carrier prepayment and postpayment reviews during March 2001. Although 100 claims were selected from each of the three carriers, 5 claims from WPS and 2 from HealthNow NY were excluded either because the billing entity did not meet our definition of physician (M.D. or D.O.) or because documentation from the carrier associated with the claim was unavailable or not interpretable.

Source: GAO analysis of independent review results.

There was slight variation in the accuracy of carrier medical review decisions for different types of payment determinations that resulted from the carriers' initial review. The independent reviewer found that carrier decisions to completely deny payment were the most accurate. In our sample, only 1 of the 64 carrier decisions (1.6 percent) to fully deny a claim was determined to be a medical review error. Carrier decisions to reduce payment amounts were slightly less accurate. The independent reviewers (with subsequent concurrence by the carriers) found errors in 5 of 59 claims (8.5 percent) that the carriers had initially decided to pay at a reduced amount. In one instance, the independent reviewer determined that the carrier should have denied the claim altogether; for the other 4 claims, it judged that the carrier should have made a smaller reduction or paid the claim in full.[20]

Three of the five instances in which the independent reviewer questioned the carrier's decision to reduce the amount paid involved claims for

[20]There was 1 claim among the 293 examined where the carrier decided (and the contractor concurred) that, based on the medical documentation provided, the physician was entitled to more than the amount submitted. In all other cases, once carrier and contractor reviews were completed, any adjustment to the claim as it was originally submitted to the carrier resulted in a decrease in the amount paid to the physician.

physicians' evaluation and management (E&M) services—commonly known as physician visits or consultations.[21] The coding system used for billing much of physician care has five separate levels of evaluation and management service intensity, each linked to a distinct payment amount. In order to assess the appropriateness of a claim's billing level, reviewers have to find specific information in the submitted clinical documentation on, among other factors, the breadth of the medical history taken, the scope of the physical examination conducted, and the complexity of the decisions made by the physician. According to CMS officials, one reason medical review decisions for these claims are likely to raise questions is that the different levels along these key dimensions are not clearly defined, such as distinguishing between "straightforward" and "low" complexity in medical decision making. Such reviews are also complicated by CMS's instruction to the carriers that they may use either the guidelines for billing evaluation and management services issued in 1995 or the ones issued in 1997, depending on which set is most advantageous to the physician.[22]

Another factor contributing to the difficulty in medically reviewing E&M claims is the broad variability in style and content found in the medical records. Carrier officials noted that some physicians meticulously document exactly what they have observed and done while others tend to be less complete and careful. Reviewers are likely to vary in what they infer from the less complete records, which, in turn, can lead to different conclusions as to whether a case is of low, medium, or high complexity.

Targeting Claims That Most Warrant Medical Review Could Be Improved

Although the carriers in our study were highly accurate in making payment determinations, they can improve their process for selecting claims for medical review that are most likely to contain billing errors. Our data show that, in fiscal year 2001, there was variation in the performance of edits—criteria used to target specific services for review—that our three carriers employed to identify medically unnecessary, or incorrectly coded, physician services. Carriers have difficulty establishing edits that routinely

[21]They can include physician encounters in hospitals and nursing facilities as well as in the doctor's office.

[22]Primary care physicians find the 1995 documentation guidelines less cumbersome, while the more detailed 1997 guidelines better reflect the needs of specialists. See Nancy-Ann DeParle, "Evaluation & Management Services Guidelines," *Journal of the American Medical Association*, vol. 283, no. 23 (June 21, 2000).

select claims with the greatest probability of errors because they have to rely, to some degree, on incomplete data. Also, CMS's oversight of the carriers does not include incentives to develop and use more refined edits. CMS has limited its involvement in this area to collecting data from the carriers on the results of reviews selected by the edits and setting general expectations for the carriers to assess the effectiveness of the edits that they use. Carriers receive no feedback on the edit effectiveness data that they have reported to the agency and little guidance as to how they could maximize the effectiveness of their procedures to select physician claims for medical review.

To help reduce local billing problems, carriers usually decide on their own which claims to select for medical review. They generally develop edits by (1) analyzing claims data to identify services or providers where local billing rates are substantially higher than national averages, and (2) selecting a small probe sample of such claims for medical review to substantiate the existence of a billing problem. Other edits are designed to ensure that physicians adhere to local medical review policies—rules that describe when and under what circumstances certain services may be covered. Claims identified by the edits are suspended, that is, temporarily held back from final processing, and the physicians involved are contacted to request the relevant medical records. Once those records arrive, claims examiners determine whether the claim should be paid in full, reduced, or denied. Of the total number of prepayment edits related to physician services used at each carrier (36 edits at WPS in each of its two largest states; 18 at NHIC's Northern California office, and 7 at HealthNow NY), 27 identified the large majority of claims undergoing medical review in fiscal year 2001. Specifically, 10 or fewer edits at each of the carriers suspended more than three-fourths of the claims medically reviewed prior to payment.

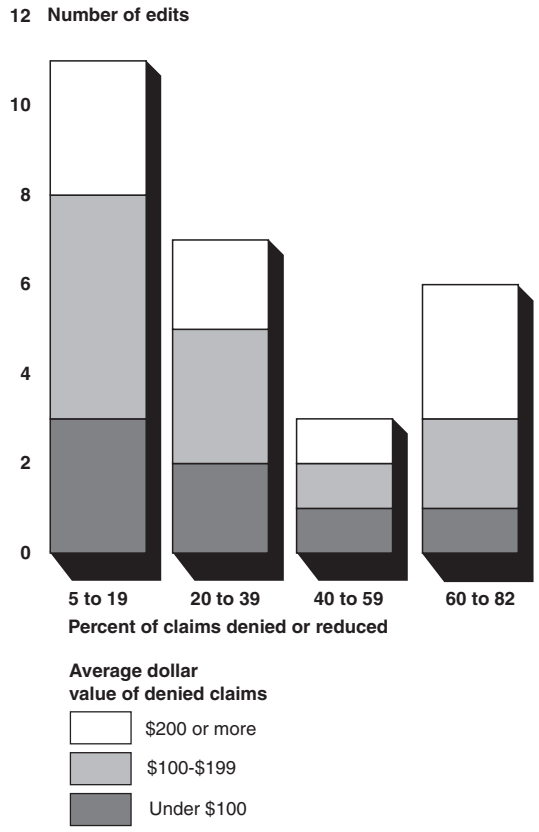
In order to assess the relative effectiveness of those edits, we drew on data that the carriers recorded on the results of reviews initiated by each edit in effect during that period. These data included information on the proportion of suspended claims that were reduced or denied as a consequence of medical review, and the average dollar reduction for those claims that were not paid in full. Edits would be considered better targeted if they have (1) a higher rate of claims denied or reduced, or (2) a larger average amount of dollars withheld from payment for an inappropriately billed service. The strongest case could be made for edits that did well on both dimensions, and the weakest case would apply for those edits that ranked low on both denial rate and average amount withheld.

Figure 1 shows the results of this analysis for the 27 prepayment edits that accounted for the largest number of claims suspended by each of our three carriers.[23] The four bars indicate the number of edits achieving different levels of denial (or reduction) rates. The grouping with the largest number of edits, 11, represents the lowest level of effective targeting, between 5 and 19 percent.[24] Two thirds of the edits, 18, have denial rates under 40 percent. By contrast, 6 edits have denial rates of between 60 and 82 percent.

[23]We have excluded some prepayment edits that are expected to have low denial rates. For instance, carriers use "pricing edits" to gather information to help determine an appropriate payment amount for newly covered services (for which Medicare has not established a set fee) or for highly complex procedures, such as certain surgeries. In these cases, although a nurse reviewer must examine each claim to determine an appropriate payment amount, claims suspended by these edits are likely to be paid in full.

[24]A once effective edit may experience declining denial rates over time, to the extent that it has its intended effect of changing physician billing behavior. This is why carriers need to monitor edit performance at periodic intervals and make appropriate adjustments.

Figure 1: Outcomes of Selected Carrier Prepayment Medical Review Edits, Fiscal Year 2001



Note: Includes data for NHIC northern California only; comparable data were not available from the carrier's southern California office.

Source: NHIC California, WPS, and HealthNow NY.

The segments within the bars indicate the average dollar amount reduced or denied when either occurs. Only 3 of the 11 major edits in the lowest denial rate group generated relatively large program savings—an average of $200 or more—for those claims that were reduced or denied. An equal number, and larger proportion, of edits in the highest denial rate group also produced savings exceeding $200 per claim.

The wide variation among these 27 major edits across both the dimensions of denial rate and average dollar amount denied or reduced suggests that there is room for improvement. CMS requires the carriers to periodically

evaluate the effectiveness of the edits they use to ensure that each has a reasonable denial rate and dollar return. However, CMS has not provided guidelines to the carriers as to how such evaluation should be conducted, or what minimum level of performance they should strive for with respect to denial rates, average dollar reductions, or other measures of efficiency. Moreover, officials at the three carriers indicated that they did not receive feedback from CMS regarding the performance of their edits, even though the carriers submit quarterly reports to the agency on the performance of their most active edits. CMS's involvement in this area was generally limited to ensuring that carriers had their own process in place for evaluating prepayment edits.

The three carriers tend to consider similar variables in evaluating edit effectiveness, but vary quite a bit in the procedures that they follow to make that assessment. In general, all three carriers consider factors such as the number of claims suspended, the denial rate, dollar savings, and the overall magnitude of the potential billing problem. With respect to process, HealthNow NY did not have any explicit procedure to evaluate edits until the end of fiscal year 2001. At that point it adopted a detailed scoring system with numeric thresholds that determine when to discontinue using a edit.[25] The other carriers continue to rely less on quantitative measures and more on the professional judgment of medical review staff in evaluating prepayment edits.

Several factors contribute to the continued use of poorly targeted edits. Some of the carriers contend that their data on the relative effectiveness of their edits are incomplete and therefore unreliable. For example, NHIC California officials noted that they often lack good information on the ultimate outcome of reviews, taking account of reversals that occur when initial carrier decisions are appealed. Not only does the appeal process take a long time, if followed to its full extent, it can also be difficult to determine why certain claim denials were overturned.[26]

[25]Each prepayment edit is scored based on its performance on six dimensions: number of claims suspended (should be greater than 150); percent of claims denied (should be greater than 25 percent); dollar value denied (should be greater than $10,000); percent of dollars denied (should be greater than 25 percent); percent of claims reversed (should be less than 40 percent); and percent of dollars reversed (should be less than 40 percent).

[26]When medical review staff denies a claim before payment, the billing physician can appeal the denial. See 42 CFR 405.801(b)(1). If the appeal is successful, the carrier may ultimately pay a claim that it initially denied. Carriers' data systems generally do not track the claims denied by medical review to determine if they are appealed and then paid.

Another reason why carriers maintain low-performing prepayment edits is that there are few incentives—and some disincentives—for them to change. In particular, carriers have agreed with CMS to conduct a certain number of reviews that are evenly distributed throughout the course of the year. Before a carrier discontinues use of an edit, it must have another one in place that will garner at least as many claims for medical review to meet workload targets, or else negotiate a change in its medical review strategy with CMS officials to reallocate those review resources to other activities. Putting new edits in place often requires carriers to adjust the selection criteria over time in order to obtain the manageable number of claims selected for review.

Carrier officials also noted that there is no systematic dissemination of carriers' best practices—those worthy of consideration by all carriers—regarding the success of individual edits or methods to evaluate edit efficiency. An official at HealthNow NY told us that they informally share information about their experiences with particular prepayment edits with other carriers operating in the same region. Carrier officials reported that this is not common practice at WPS or NHIC California. In a 1996 report on selected prepayment edits, we recommended that HCFA, now CMS, disseminate information to carriers on highly productive edits.[27] However, the agency currently does not identify and publicize in any systematic manner those edits that generate high denial rates or the selection criteria used to develop them.

[27]U.S. General Accounting Office, *Medicare: Millions Can Be Saved by Screening Claims for Overused Services*, GAO/HEHS-96-49 (Washington, D.C.: January 30, 1996).

CMS Makes Claims Accuracy New Benchmark for Measuring Carrier Performance

Since 1996, the overall level of payment errors for the Medicare program has been tracked nationwide in annual audit reports issued by the HHS Office of Inspector General (OIG). In the most recent audit, covering fiscal year 2001, the OIG found that $12.1 billion, or about 6.3 percent of the $191.8 billion in processed fee-for-service payments, was improperly paid to Medicare providers.[28] These OIG reports of aggregate Medicare payment errors have spurred CMS to improve its efforts to safeguard Medicare payments by assessing not only an error rate nationwide but also for the individual carriers.

In February 2000, HCFA announced the development of a new tool to assess individual carrier performance called the Comprehensive Error Rate Testing (CERT) Program. CERT is designed to measure, for all claims, the accuracy of payment decisions made by each carrier.[29] The CERT benchmark will allow CMS to hold the carriers accountable for the accuracy of payment decisions for all claims processed, not just those selected for review. Thus, the results will reflect not only the carrier's performance, but also the billing practices of the providers in their region. According to CMS officials, CERT information on all the carriers processing physician claims is expected to become available in November 2002. At that point, both CMS and the carriers can begin to use that information for program oversight and management, and will then see if the expectations for CERT are met in practice.

Under the CERT program, CMS will use an independent contractor to select a random sample of approximately 200 claims for each carrier from among all those submitted each month for processing. For this sample, the carrier will provide the CERT contractor with information on the payment decisions made and all applicable medical documentation used in any medical reviews of the sample claims. The CERT contractor will request comparable documentation from physicians whose claims in the sample

[28]For the fiscal year 2001 audit, OIG selected 600 beneficiaries nationwide with 6,594 fee-for-service claims processed for payment. Based on this sample, it estimated the range of improper payments at the 95 percent confidence level to be $7.2 billion to $16.9 billion. The OIG indicated that this result was not significantly different from the estimates for the past 3 years. See Department of Health and Human Services/Office of Inspector General, *Improper Fiscal Year 2001 Medicare Fee-For-Service Payments*, A-17-00-02000 (Washington, D.C.: Feb. 15, 2002).

[29]Previously, carriers selected a portion of their prepayment medical reviews through a random sampling procedure. CERT is taking the place of that random sample, and henceforth carriers should only select claims for prepayment review that have been identified as potentially problematic.

were not medically reviewed by the carrier. The CERT teams of clinician reviewers will examine the documentation and apply the applicable national and local medical policies to arrive at their own payment decisions for all of the sampled claims.

With the development of carrier-level error rates, CMS expects to monitor payment accuracy trends for the individual carriers and focus its oversight on those carriers with relatively high, or worsening, rates of error. Moreover, on a national basis, CERT will calculate error rates for different provider types. For example, it will indicate how often physicians bill incorrectly and receive either too much or too little payment compared to such nonphysician providers as ambulance companies and clinical labs. The structure of subgroup analyses designed to help carriers better target their medical reviews remains open to discussion among CMS officials.

CERT will complement but not replace CMS tracking systems designed to monitor carrier performance using data periodically reported to CMS by the carriers concerning medical review costs, the reduction in provider payments resulting from medical reviews, and workload. CMS has relied on these data to ensure that carriers sustain the level of effort specified in agreements with CMS—particularly the number of medical reviews conducted. CMS is currently working to consolidate and streamline these various reports into a Program Integrity Management Reporting (PIMR) system. CMS's intention is for PIMR to collect, from each carrier, data such as the number of claims medically reviewed, the number denied, the number of denials reversed on appeal, and the associated dollar amounts saved or recouped. Currently, this information is not maintained in a common format and is difficult to compile. The first management reports based on PIMR are expected by the end of fiscal year 2002.

In addition to CERT and the carrier-reported data, CMS oversight of physician medical review will continue to rely on contractor performance evaluations (CPEs)—assessments based on site visits conducted by a small team of CMS regional and headquarters staff. For carrier medical review activities, these CMS evaluations occur at irregular intervals, depending on the carrier's volume of claims and the level of risk of finding substantial problems at the carrier. CMS's evaluation emphasizes an assessment of the carrier's compliance with Medicare rules and procedures in areas related to medical review—such as data analysis to support the selection of edits, the development of local coverage rules,

and tracking contacts with physicians. The evaluation also involves examining a small number of claims to determine the accuracy of the carrier's review decisions.[30] Critics have previously alleged that CPE assessments lacked consistency and objectivity. In response, CMS has attempted to ensure greater uniformity across carriers in the way these evaluations are conducted by recruiting CPE team members from the agency as a whole, not the local regional office, and by using nationally based CPE protocols.

Concluding Observations

While CMS has modified its medical review procedures, it is too soon to determine whether the PCA approach will enhance the agency's efforts to perform its program integrity responsibilities. Carrier staff conduct medical reviews to maintain program surveillance and make physicians aware of any billing practices that are not in keeping with payment rules. In this regard, CMS's PCA policy emphasizes feedback and educational contacts with individual physicians. Evaluating the efficacy of this policy will require a systematic examination of carriers' performance data. When CERT data become available, CMS may be in a better position to assess PCA's impact on reducing billing errors and preventing inappropriate payments.

Agency Comments

CMS officials reviewed a draft of this report and generally agreed with its findings. In particular, the agency noted that our discussion of the effectiveness of carrier edits confirmed the need for CMS to "become more active in assisting contractors in this area." The agency also provided a number of technical corrections and clarifications that we incorporated into the text as appropriate. These comments are reprinted in Appendix III.

We are sending copies of this report to the Administrator of CMS and we will make copies available to others upon request. In addition, the report will be available at no charge on the GAO Web site at http://www.gao.gov.

[30]In a typical CPE about 30 claims are reassessed. By contrast, CERT will examine approximately 200 claims payment decisions per month for each carrier.

If you or your staffs have questions about this report, please contact me at (312) 220-7600 or Rosamond Katz at (202) 512-7148. Other contributors to this report were Hannah Fein, Jenny Grover, Joel Hamilton, and Eric Peterson.

Leslie G. Aronovitz,
Director, Health Care—Program
Administration and Integrity Issues

List of Committees

The Honorable Max Baucus
Chairman
The Honorable Charles E. Grassley
Ranking Minority Member
Committee on Finance
United States Senate

The Honorable W.J. "Billy" Tauzin
Chairman
The Honorable John D. Dingell
Ranking Minority Member
Committee on Energy and Commerce
House of Representatives

The Honorable William M. Thomas
Chairman
The Honorable Charles B. Rangel
Ranking Minority Member
Committee on Ways and Means
House of Representatives

Appendix I: Review of Medicare Carrier Medical Review Decisions

We assessed the claims review accuracy of the three carriers in our study—National Heritage Insurance Company in California, Wisconsin Physicians Service Insurance Corp, and HealthNow NY—by validating initial medical review decisions involving physician claims. We contracted with DynCorp—the Medicare contractor already selected by CMS to administer its Comprehensive Error Rate Testing (CERT) program—to use the same review procedures developed for CERT in assessing a sample of medical review decisions made by the three carriers.

Methodology Used to Validate Carrier Medical Review Decisions

We requested that each carrier identify the universe of physician claims subjected to prepayment and postpayment review during March 2001, limiting the universe to those claims submitted by M.D.s and D.O.s. From that universe, Dyncorp randomly selected 100 claims for review. Then, DynCorp obtained the medical record information for those claims from the carrier, and reviewed each payment decision for accuracy. The number of carrier decisions examined by DynCorp staff exceeded the number of claims because, in several instances, carriers had reviewed multiple lines on a claim. The results of this assessment of carrier medical review decisions can only be generalized to the universe of claims from which the samples were drawn: claims from M.D.s or D.O.s that underwent medical review in March 2001 by one of our three carriers.

In reviewing payment accuracy, DynCorp staff was tasked with determining if the carrier's initial review decision was supported by the medical record and carrier policies in place at the time the payment decision was made. Specifically, DynCorp assessed whether documentation in the medical records supported the procedure codes and level of service that was billed. Where their determination differed from that of the carrier, DynCorp staff discussed those claims with the carrier's medical review staff. In all but one case, the parties came to agreement on whether payment decisions were accurate. In the one case where agreement could not be reached, the acting deputy director of CMS's Program Integrity Group—a physician—provided a second opinion that confirmed the carrier's decision.

Results of the DynCorp Review

The results obtained from DynCorp's review of physician claims undergoing medical review were consistent across the three carriers. The accuracy of decisions across all the sampled medical reviews for each carrier exceeded 94 percent. (See table 7.) In those cases where medical review errors were identified, NHIC California and WPS decisions resulted in a mix of underpayments and overpayments. However, HealthNow NY's

review errors were concentrated in decisions to pay claims in full that should have been denied or reduced.

Table 7: Accuracy of Medical Review Decisions on Physician Claims by Carrier

		Inaccurate decision rate	
Carrier decision	Accurate decision rate (percent)	Overpayment (percent)	Underpayment (percent)
All decisions on sampled claims			
NHIC California[a] (n=100)	94.0	3.0	3.0
WPS[b] (n=95)	96.8	1.1	2.1
HealthNow NY[b] (n=98)	95.9	4.1	0.0
Deny payment			
NHIC California (n=24)	95.8	0.0	4.2
WPS (n=26)	100.0	0.0	0.0
HealthNow NY (n=14)	100.0	0.0	0.0
Pay in part			
NHIC California (n=26)	88.5	3.8	7.7
WPS (n=20)	90.0	0.0	10.0
HealthNow NY (n=13)	100.0	0.0	0.0
Pay in full			
NHIC California[a] (n=50)	96.0	4.0	0.0
WPS (n=49)	98.0	2.0	0.0
HealthNow NY[c] (n=71)	94.4	5.6	0.0

[a]Includes six claims in the NHIC California sample that DynCorp (with concurrence of the acting deputy director of CMS's Program Integrity Group) judged inaccurate on technical grounds, but we considered them to have been decided appropriately. Although DynCorp noted that the physicians had provided documentation sufficient to justify payment, it judged them to have been paid in error because NHIC California had a local medical review policy (LMRP) that explicitly required the physician to document the number of minutes spent with the patient for these services. The physicians submitting these claims had not done so, but the carrier and DynCorp agreed that if the LMRP had not been in place, the documentation provided would have been sufficient to justify payment. Therefore, the problem identified by the DynCorp review was the failure of NHIC California to update its LMRP to reflect the review practices it was actually following—not inadequacies in the reviews themselves.

[b]Although 100 claims were selected from each of the three carriers, 5 claims from WPS and 2 from HealthNow NY were excluded either because the billing entity did not meet our definition of physician (M.D. or D.O.) or because documentation from the carrier associated with the claim was unavailable or not interpretable.

[c]Includes one claim where the carrier determined that the physician was entitled to more than the amount submitted.

Source: GAO analysis of claims review data from NHIC California, HealthNow NY, WPS, and Dyncorp.

Because a relatively small proportion of medical reviews are conducted after claims payment, our samples from the three carriers included just 19 claims where a postpayment review was performed. The accuracy of carrier determinations for both prepayment and postpayment medical reviews was consistent, at about 95 percent. (See table 8.)

Table 8: Accuracy of Prepayment and Postpayment Medical Review Decisions on Physician Claims (percent)

		Inaccurate decision rate	
Medical review	Accurate decision rate (percent)	Overpayment (percent)	Underpayment (percent)
Prepayment (n=274)	95.6	2.9	1.5
Postpayment (n=19)	94.7	0.0	5.3

Source: GAO analysis of claims review data from NHIC California, HealthNow NY, WPS, and Dyncorp.

Appendix II: Recovery of Overpayments

Carriers attempt to collect any overpayments due the Medicare program as soon as possible after the completion of postpayment reviews. The carrier notifies physician practices that they have three options for returning an overpayment: (1) pay the entire overpayment amount within 30 days, (2) apply for an extended repayment plan, or (3) allow the carrier to offset the overpayment amount against future claims.

Initially, the carrier sends a letter informing the physician practice of the medical review results and the specific dollar amount that the practice must return to Medicare. The letter provides an explanation of the procedures for repaying an overpayment, which includes a statement of Medicare's right to recover overpayments and charge interest on debts not repaid within 30 days, as well as the practice's right to request an extended repayment plan if the overpayment cannot be paid in that time.[1] The letter also advises the physician practice of the right to submit a rebuttal statement prior to any recoupment by the carrier and to appeal the review decision to, in the first instance, the carrier's separate appeals unit. In addition, the letter notifies the practice of any additional reviews that the carrier has planned. Regardless of whether the physician practice appeals the review decision, repayment is due within 30 days of the date of the letter, unless an extension is approved.

Carriers will consider extended repayment plans for those physician practices that cannot make a lump sum payment by the due date. To qualify for an extension, the overpayment amount must be $1,000 or more and a practice must prove that returning an overpayment within the required time period would cause a financial hardship. Accordingly, a physician practice must offer specific documentation to support the request, including a financial statement with information on monthly income and expenses, investments, property owned, loans payable, and other assets and liabilities. In addition, if the requested repayment extension is for 12 months or longer, the physician practice must submit at least two letters from separate institutions indicating that they denied a loan request for the amount of the repayment. Requests for payment extensions that exceed 12 months must be referred to CMS regional staff for approval.

[1]Medicare regulations provide for the assessment of interest at the higher of the private consumer rate or the current value of funds rate (5 percent for calendar year 2002). See 42 C.F.R. Sec. 405.378 (d) (2001). As of February 1, 2002, the private consumer rate was 12.625 percent.

If a physician practice does not return payment within 30 days or establish a repayment extension plan, the carrier must offset the amount owed against pending or future claims. The carrier has some discretion as to the exact date that offsetting begins, taking into consideration any statements or evidence from the physician practice as to the reasons why offsetting should not occur.[2] In fiscal year 2001, HealthNow NY offset amounts owed by 72 of 95 physician practices that did not pay their overpayment amounts within 30 days. Most of the practices that did not have amounts offset returned their overpayments within 40 days. Any offset payments are applied against the accrued interest first, and then the principal.

As shown in table 9, the three carriers in our study reported that most physician practices assessed an overpayment in fiscal year 2000 or 2001 repaid Medicare within 6 months of the carrier's notice.

[2]CMS instructs carriers to allow physician practices 15 days from the initial notification letter to submit information related to offsetting. The carrier is to promptly consider and respond to the information. If the carrier does not receive such a response from the practice, CMS instructs carriers to initiate offset within 40 days after the date of the letter notifying the physician practice of the overpayment amount.

Table 9: Recovery of Overpayments From Physician Practices, Fiscal Years 2000-2001

	NHIC California[a]		WPS		HealthNow NY	
	2000	2001	2000	2001	2000	2001
Number of practices assessed an overpayment[b]	58	81	106[c]	76[c]	158	151
Number of practices that repaid within:						
30 days	33	48	37	24	46	56
31-180 days	23	33	57	34	107	94
181-365 days	0	0	1	1	3	1
Over 1 year	2	0	6	4	1	0
Outstanding	0	0	5	13	1	0
Number of practices with overpayments of $5,000 or more that repaid within:						
30 days	21	10	7	2	1	3
31-180 days	12	17	24	13	6	1
181-365 days	0	0	1	0	1	0
Over 1 year	2	0	6	1	1	0
Outstanding	0	0	5	2	0	0

[a]Data include NHIC's northern California office only; data were not available for its southern office for fiscal year 2000. However, during fiscal year 2001, the southern office assessed 137 physician practices an overpayment, and all but 5 repaid within 6 months.

[b]Some overpayment assessments may reflect the outcomes of medical reviews conducted in the previous fiscal year.

[c]WPS data represent the number of overpayment assessments. Because a few physicians practices were assessed more than one overpayment during the fiscal year, these data very slightly overstate the number of practices involved.

Source: NHIC California, WPS, and HealthNow NY.

The three carriers also reported few requests from physician practices for extended repayment plans. As shown in table 10, none of the carriers had more than four requests during fiscal year 2001, and no extension exceeded 1 year.

Table 10: Requests for Repayment Extensions, Fiscal Years 2000-2001[a]

	NHIC California		WPS		HealthNow NY	
	2000	2001	2000	2001	2000	2001
Extensions requested	12	4	3	1	0	1
Extensions approved	11	2	3	0	[b]	1
Extension period granted	6 to 24 months	6 to 12 months	7 to 12 months	[b]	[b]	6 months
Range of overpayments[c] (dollars)						
Lowest amount	$18,337	$20,588	$2,708	[b]	[b]	[b]
Median amount	159,894	[b]	28,749	[b]	[b]	$13,343
Highest amount	324,106	49,981	105,924	[b]	[b]	[b]

[a]Some requests for repayment extensions relate to overpayments assessed in a previous fiscal year.

[b]Not applicable.

[c]Repayment amounts include the principal only, adjusted for any reductions that may have resulted from physician rebuttals or appeals.

Source: NHIC California, WPS, and HealthNow NY.

Appendix III: Comments from the Centers for Medicare and Medicaid Services

DEPARTMENT OF HEALTH & HUMAN SERVICES

Centers for Medicare & Medicaid Services

Administrator
Washington, DC 20201

MAY 15 2002

TO: Leslie G. Aronovitz
Director, Health Care—Program
Administration and Integrity Issues
General Accounting Office

FROM: Thomas A. Scully
Administrator
Centers for Medicare & Medicaid Services

SUBJECT: General Accounting Office (GAO) Draft Report, *Medicare: Recent CMS Reforms Address Carrier Scrutiny of Physicians' Claims for Payment* (GAO-02-693)

Thank you for the opportunity to review the above-referenced GAO draft report.

We are appreciative that GAO finds our program integrity efforts to improve medical review procedures have had a positive effect on physician oversight. The findings you report were certainly some of the goals we worked towards when implementing these program changes. We plan to continue and increase provider education with the goal of reducing the payment error rate.

The Comprehensive Error Rate Testing (CERT) program will calculate three rates: 1) a paid claims error rate; 2) a claims processing error rate; and 3) a provider compliance rate. The program will calculate an error rate for four levels—national, contractor specific, benefit specific, and provider type. This program will facilitate the Centers for Medicare & Medicaid Services' ability to take appropriate corrective actions and can be used to better manage contractor performance. The Progressive Corrective Action (PCA) initiative is designed to focus resources and corrective actions on those areas where there is the greatest risk of inappropriate payments.

We appreciate your discussion of edit effectiveness. It is a subject that we have discussed often internally over the past year, and your comments confirm our belief that we need to become more active in assisting contractors in this area.

We have attached a number of technical comments that we believe will better characterize our progressive corrective action and the Comprehensive Error Rate Testing activities. We look forward to working with GAO on this and other issues.

Attachment

Related GAO Products

Medicare: Communications With Physicians Can Be Improved (GAO-02-249, February 27, 2002).

Medicare Management: CMS Faces Challenges to Sustain Progress and Address Weaknesses (GAO-01-817, July 31, 2001).

Medicare Management: CMS Faces Challenges in Safeguarding Payments While Addressing Provider Needs (GAO-01-1014T, July 26, 2001).

Regulatory Issues for Medicare Providers (GAO-01-802R, June 11, 2001).

GAO's Mission

The General Accounting Office, the investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.

Obtaining Copies of GAO Reports and Testimony

The fastest and easiest way to obtain copies of GAO documents at no cost is through the Internet. GAO's Web site (www.gao.gov) contains abstracts and full-text files of current reports and testimony and an expanding archive of older products. The Web site features a search engine to help you locate documents using key words and phrases. You can print these documents in their entirety, including charts and other graphics.

Each day, GAO issues a list of newly released reports, testimony, and correspondence. GAO posts this list, known as "Today's Reports," on its Web site daily. The list contains links to the full-text document files. To have GAO e-mail this list to you every afternoon, go to www.gao.gov and select "Subscribe to daily E-mail alert for newly released products" under the GAO Reports heading.

Order by Mail or Phone

The first copy of each printed report is free. Additional copies are $2 each. A check or money order should be made out to the Superintendent of Documents. GAO also accepts VISA and Mastercard. Orders for 100 or more copies mailed to a single address are discounted 25 percent. Orders should be sent to:

U.S. General Accounting Office
441 G Street NW, Room LM
Washington, D.C. 20548

To order by Phone:
Voice: (202) 512-6000
TDD: (202) 512-2537
Fax: (202) 512-6061

To Report Fraud, Waste, and Abuse in Federal Programs

Contact:

Web site: www.gao.gov/fraudnet/fraudnet.htm
E-mail: fraudnet@gao.gov
Automated answering system: (800) 424-5454 or (202) 512-7470

Public Affairs

Jeff Nelligan, managing director, NelliganJ@gao.gov (202) 512-4800
U.S. General Accounting Office, 441 G Street NW, Room 7149
Washington, D.C. 20548