

Audit of 1996 Medicare HEDIS Data

FINDINGS

Prepared for:

Health Care Financing Administration

Prepared by:

I PRO

Managed Care Department
1979 Marcus Avenue, First Floor
Lake Success, NY 11042

Table of Contents

EXECUTIVE SUMMARY2

CHAPTER 1: BACKGROUND AND METHODOLOGY9

CHAPTER 2: AUDIT FINDINGS—SELECTED HEDIS MEASURES15

CHAPTER 3: AUDIT FINDINGS — INFORMATION SYSTEMS AND PROCESSES22

CHAPTER 4: CONCLUSIONS28

CHAPTER 5: RECOMMENDATIONS.....31

EXECUTIVE SUMMARY

Introduction

Purpose

The audit of 1996 Medicare HEDIS[®] 3.0 data had four main purposes:

1. Determine, based upon a sample of health plans, if the 1996 Medicare HEDIS data submitted to HCFA were prepared in compliance with HEDIS specifications.
2. Determine, based upon a sample of health plans, if the 1996 Medicare HEDIS data are usable for health plan comparison.
3. Test certain auditing methods and help determine future Medicare HEDIS reporting and validation requirements.
4. Collect information for HCFA to test and improve HEDIS specifications.

Background

In Operational Policy Letter #47, HCFA mandated that all Section 1876 Risk or Cost managed care plans report calendar year 1996 data to HCFA on 32 selected HEDIS 3.0 measures and participate in an audit of their information systems and HEDIS-reporting processes. For purposes of reporting HEDIS measures, several Medicare contracts that cover multiple metropolitan areas were divided into markets. When the markets were tabulated with contracts, the final number of contract-markets (i.e., the reporting unit for HEDIS) that reported the required HEDIS measures totaled 284.

IPRO, the New York State PRO, was selected as the prime contractor to lead the audit of 1996 Medicare HEDIS 3.0 data. IPRO further subcontracted with Stratis Health, the Minnesota PRO, and MetaStar, the Wisconsin PRO, to assist in conducting audit site visits. In addition, The MEDSTAT Group was selected to provide technical assistance and additional staffing to IPRO.

Of the 32 HEDIS 3.0 measures required for reporting to HCFA, five measures were selected for audit. Four Effectiveness of Care measures were selected:

- Breast Cancer Screening,
- Eye Exams for People with Diabetes,
- Beta Blocker Treatment After a Heart Attack, and
- Follow-up After
- Hospitalization for Mental Illness.

The fifth measure, Frequency of Selected Procedures, was selected from the Use of Services Domain, and included 10 procedures.

HEDIS[®] is a registered trademark of the National Committee for Quality Assurance (NCQA).

Methodology

The audit methodology was designed by IPRO in consultation with a Technical Advisory Panel, which consisted of representatives from HCFA, IPRO, NCQA, The MEDSTAT Group, several health plans, and other PROs. The audit consisted of two phases: a desk audit and an onsite audit.

For the desk audit, all health plans that submitted 1996 HEDIS 3.0 Medicare data were mailed a Baseline Assessment to complete. Health plans were required to provide information about their membership, data systems, provider networks, ancillary services (including vendors), and methods and processes used to prepare and submit rates for required Medicare HEDIS measures. All 284 contract-markets returned a completed Baseline Assessment.

A sample of 79 contract-markets was selected for an onsite audit using the Probability Proportional to Size (PPS) sampling method. The PPS method weights more heavily those contracts with large enrollments, while still allowing for representation of contracts with smaller enrollments. The selected sample represented 65 percent of the December 1996 Medicare managed care beneficiary population.

Onsite audits were conducted to follow-up on issues identified in the Baseline Assessment and to more fully determine a health plan's ability to accurately prepare HEDIS measures. As part of the onsite visit, auditors interviewed staff responsible for overseeing health plan processes that affect HEDIS reporting, including claims/encounter processing, medical record review, membership data processing, provider data processing, data integration, and HEDIS-specific computer programming. Auditors confirmed information provided in documents and interviews through live demonstrations of key health plan systems.

The onsite audit team consisted of 2 to 5 individuals with diverse backgrounds, including health care data analysts, programmers, statisticians, systems analysts, and medical record review supervisors. To ensure inter-auditor consistency, all auditors attended a rigorous, four-day training session and used a standard set of data collection tools developed by IPRO. Auditors followed the 1997 NCQA HEDIS Compliance Audit[™] Standards and Guidelines during their onsite reviews of health plan information systems and HEDIS preparation processes.

An onsite audit visit typically lasted from 2 to 5 days. After the site visits, health plans were given Final Reports, which provided narrative descriptions of audit findings and a final score—or “audit designation”—for each of the five audited measures.

Study Limitations

In reviewing the audit findings, it is important to note the limitations of the study, which include:

- Only five HEDIS measures, from only two measure domains, were audited and, thus, the audit findings could not be generalized to all HCFA-required measures.

[™] NCQA HEDIS Compliance Audit is a trademark of the National Committee for Quality Assurance (NCQA).

- Onsite audits were performed only on a sample of primarily large health plans and, as a number of reporting problems are plan-specific, the audit findings could not be generalized to all Medicare health plans.
- While health plan processes for medical record review were audited, a re-review of a sample of medical records was not part of the audit.
- While health plan claims/encounter forms and clinical coding systems were evaluated, an assessment of how accurately providers document clinical information on the required claims/encounter forms was not included in the audit.

Audit Findings: Selected HEDIS Measures

Effectiveness of Care Measures

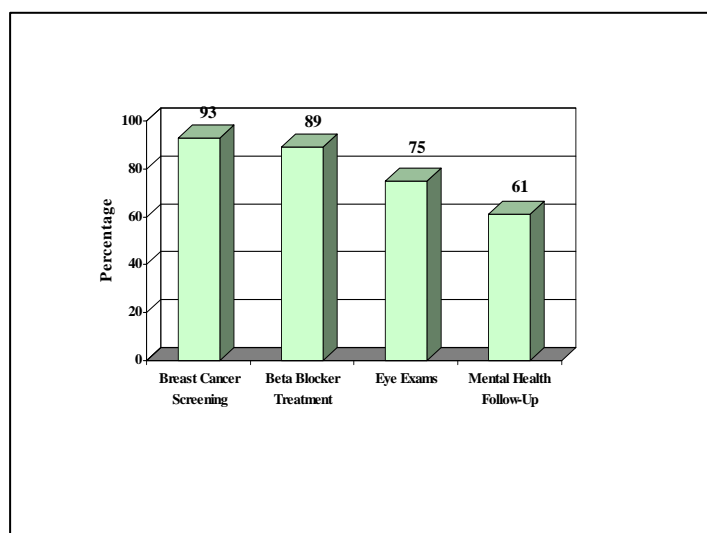
Table 1 summarizes the audit designations earned for the four Effectiveness of Care measures that were reviewed during the onsite audits for 79 contract-markets. The audit designations were assigned using the scoring system specified in the 1997 NCQA HEDIS Compliance Audit Standards and Guidelines. An audit designation of “Report” (R) indicates that the measure was fully or substantially compliant with HEDIS specifications. A designation of “Not Report” (NR) indicates that the measure was not in compliance with HEDIS specifications or that the health plan chose not to report a measure even though reporting was required. If the auditors determined that the sum of the errors made by the health plan caused the reported rate to deviate by more than 5 percentage points from the true rate—or if the health plan could not produce information to document the extent of the deviation—then the rate was designated as non-compliant with HEDIS specifications and given an “NR” designation. If the deviation was less than 5 percentage points, then the rate was considered compliant and given an “R” designation. The audit designation of “Not Applicable” (NA) indicates that the health plan did not have sufficient population to report a rate.

Table 1 Audit Designations using the NCQA Scoring System

| Effectiveness of Care Measures | R | NR | NA |
|--|-------|-------|-------|
| Breast Cancer Screening | 88.6% | 6.3% | 5.1% |
| Beta Blocker Treatment After a Heart Attack | 69.6% | 8.9% | 21.5% |
| Eye Exams for People with Diabetes | 73.4% | 24.1% | 2.5% |
| Follow-up After Hospitalization for Mental Illness | 45.6% | 29.1% | 25.3% |

Another way to evaluate the audit findings is to exclude contract-markets with designations of NA and then recalculate the percentage of health plans that accurately produced a rate. Below, Figure 1 illustrates the percentage of contract-markets that reported rates that were fully or substantially compliant with HEDIS specifications, based only on those contract-markets that were required to report a rate for a measure. Specifically, the percentages shown in the figure below were calculated by dividing the number of contract-markets with designations of R by the number of contract-markets with designations of either R or NR, with NA excluded from the calculation.

Figure 1 Percentage of Contract-Markets with Fully or Substantially Compliant HEDIS Data Based Only on Health Plans With Samples Adequate to Report the Measure



Among health plans that attempted to produce the Effectiveness of Care measures, a designation of Not Report was most often issued because the health plan reported a rate that was more than 5 percentage points above the true rate; the magnitude of over-reporting was generally between 5 to 12 percentage points. About one third of the Not Report designations occurred because a health plan could not document the magnitude of an error's effect on the reported rate. In only one instance (representing 2 percent of the Not Report designations) did a health plan receive a Not Report designation for reporting a rate that was more than 5 percent below the true rate.

Frequency of Selected Procedures

Unlike the rates for the Effectiveness of Care measures that can be calculated using the hybrid method, which allows for better data capture through medical record review, the Frequency of Selected Procedures measure must be calculated using administrative data only. Consequently, rates for this measure are more susceptible to underreporting. During the onsite audit visits, it became evident that some health plans used substantially incomplete databases to report this measure.

As the 1997 NCQA HEDIS Compliance Audit Standards and Guidelines did not explicitly specify that the use of incomplete data is sufficient cause to “fail” a measure, this report profiles the audit findings for the Frequency of Selected Procedures in two different formats. Table 2a summarizes the audit designations for each procedure when the use of incomplete data is not considered as a significant deviation. Table 2b shows how the percentages are altered when the true impact of incomplete data is included in determining the audit designations.

**Table 2a: Compliance Scores—
Frequency of Selected Procedures
(Effect of incomplete data not considered)**

| <i>Procedure</i> | R | NR | NA |
|----------------------------|-----|-----|----|
| CABG | 90% | 10% | 0% |
| Angioplasty (PTCA) | 96% | 4% | 0% |
| Carotid Endarterectomy | 98% | 2% | 0% |
| Fracture of Femur | 95% | 5% | 0% |
| Total Hip Replacement | 98% | 2% | 0% |
| Total Knee Replacement | 98% | 2% | 0% |
| Pt. Excision of Lg Intest. | 91% | 9% | 0% |
| Cholecystectomy, Open | 96% | 4% | 0% |
| Cholecystectomy, Closed | 96% | 4% | 0% |
| Hysterectomy | 96% | 4% | 0% |
| Prostatectomy | 96% | 4% | 0% |

**Table 2b: Compliance Scores—
Frequency of Selected Procedures
(Effect of incomplete data considered)**

| <i>Procedure</i> | R | NR | NA |
|----------------------------|-----|-----|----|
| CABG | 71% | 29% | 0% |
| Angioplasty (PTCA) | 72% | 28% | 0% |
| Carotid Endarterectomy | 74% | 26% | 0% |
| Fracture of Femur | 71% | 29% | 0% |
| Total Hip Replacement | 74% | 26% | 0% |
| Total Knee Replacement | 74% | 26% | 0% |
| Pt. Excision of Lg Intest. | 72% | 28% | 0% |
| Cholecystectomy, Open | 74% | 26% | 0% |
| Cholecystectomy, Closed | 74% | 26% | 0% |
| Hysterectomy | 74% | 26% | 0% |
| Prostatectomy | 74% | 26% | 0% |

As indicated in Table 2a, contract-markets used measure-preparation processes that significantly deviated from HEDIS specifications less than 10 percent of the time for all eleven of the procedures and less than 5 percent of the time for all but two of the procedures. Yet, when the effect of incomplete data is considered, the percentage of “NR” designations increase, on average, from 4 percent to 27 percent. For all procedures, when the effect of incomplete data is considered, less than 75 percent of contract-markets reported a fully or substantially compliant rate.

The magnitude of under-reporting because of incomplete data varied tremendously. In many cases, it was clear that incomplete data caused under-reporting of more than 5 percent, but the health plan had little or no documentation for auditors to use to estimate the actual magnitude. When data were available, the effect of incomplete data generally ranged from 10 percent to 40 percent or more.

Audit Findings: Information Systems and Processes

As part of their review of health plans' information systems and HEDIS preparation processes, auditors focused on several key areas, as outlined in the 1997 NCQA HEDIS Compliance Audit Standards and Guidelines. While many health plans have sound information systems and HEDIS preparation processes, many other health plans had one or more significant information system or process limitations. A summary of the most common challenges to HEDIS reporting are identified in each area as follows:

Medical Data—Forms and Coding: The most common concerns were the use of proprietary forms that capture incomplete diagnosis and procedure information and the use of internally developed coding systems that were difficult to accurately translate into the standard coding (e.g., ICD-9, CPT-4) used in HEDIS specifications.

Medical Data—Data Transfer and Entry: The most frequent problems were poor data capture, inadequate or incompatible information systems, inefficient or poorly monitored data collection and processing procedures, and insufficient oversight of vendors.

Medical Record Review: The most common problems were the lack of training and oversight of medical record reviewers, the use of data-collection instruments that provided little guidance to reviewers to ensure that they captured accurate information, and the reliance upon provider attestation in place of medical record review (particularly for the Follow-Up After Hospitalization for Mental Illness measure).

Membership Data: Some health plans had difficulty tracking members as they changed status or moved from one product line to another, while others did not maintain enough member history to fully support continuous enrollment calculations for all measures.

Provider Data: The onsite audit found that key provider information—such as specialty, credentials, tax identification numbers, and locations of service—were sometimes inaccurate, incomplete, or maintained on multiple information systems that could not be linked with each other or with claims/encounter and membership databases.

Data Integration and Control: Key concerns included insufficient process documentation, inadequate information system staffing, system integration problems, and failures to follow standard information system protocols.

Documentation: Documentation of HEDIS preparation was often inadequate, particularly for computer programs, medical record review processes, vendor data, and mappings of proprietary codes to standard codes.

It is important to note that the information system problems listed above sometimes caused only minor deviations in reported HEDIS rates but at other times caused significant deviations. Whether a particular problem caused a minor or significant deviation depended upon the systems and processes used by a particular plan, the requirements of the particular HEDIS measure, and the presence or absence of other problems that could have either a reinforcing or mitigating effect.

Conclusions/Recommendations

IPro's conclusions/recommendations regarding the findings and release of data for the five Medicare HEDIS measures are as follows:

Breast Cancer Screening and Beta Blocker Treatment After A Heart Attack

As indicated earlier in Figure 1, these measures were found to be prepared in full or substantial compliance with HEDIS specifications by 93 and 89 percent, respectively, of the audited contract-markets that reported a rate. If HCFA deems this level of accuracy to be adequate and, therefore, chooses to publicly release comparisons of health plan performance on these two indicators, IPro recommends that a clear statement of data limitations be included. Such a statement is particularly important because 205 of the 284 contract-markets did not undergo an onsite audit and, consequently, some biased rates would be reported. In addition, the accuracy of medical record review was not validated for the approximately 60 percent of health plans that used the hybrid method to prepare these measures.

Eye Exams for People with Diabetes and Follow-Up After Hospitalization for Mental Illness

Less than 80 percent of the contract-markets in the audit sample prepared the Eye Exams for People with Diabetes and Follow-up After Hospitalization for Mental Illness measures in full or substantial compliance with HEDIS specifications. Due to the high rate of failure and the numerous difficulties experienced by some health plans in reporting these two measures, IPro does not recommend that HCFA use these measures for health plan comparison.

Frequency of Selected Procedures

IPro also recommends that HCFA not use the Frequency of Selected Procedures measure for health plan comparison. While the processes used to prepare this measure were found to be compliant with HEDIS specifications for over 90 percent of the contract-markets in the audit sample, rates for approximately 23 percent of the contract-markets were prepared using databases that were substantially incomplete. Consequently, because of a large variety of information systems and process issues, some health plans' reported rates understate true performance.

Unaudited Measures

Based upon our findings for the five audited measures—as well as our findings regarding health plan information systems and HEDIS preparation processes—IPro recommends that HCFA be extremely cautious in using any unaudited measures for health plan comparison.

Additional Recommendations

In Chapter 5 of the full report, we provide recommendations to health plans, NCQA, and HCFA regarding process changes that could help improve the accuracy, completeness, and timeliness of Medicare HEDIS reporting. The chapter also includes suggestions for increasing the effectiveness of future Medicare HEDIS audits.

CHAPTER 1: BACKGROUND AND METHODOLOGY

Purpose

The audit of 1996 Medicare HEDIS 3.0 data had four main purposes:

1. Determine, based upon a sample of health plans, if the 1996 Medicare HEDIS data submitted to HCFA were prepared in compliance with HEDIS specifications.
2. Determine, based upon a sample of health plans, if the 1996 Medicare HEDIS data are usable for health plan comparison.
3. Test certain auditing methods and help determine future Medicare HEDIS reporting and validation requirements.
4. Collect information for HCFA to test and improve HEDIS specifications.

Background

In December of 1996, HCFA distributed Operational Policy Letter #47, titled New Requirements for Medicare Health Plans in 1997: HEDIS 3.0 Measures and the Medicare Beneficiary Satisfaction Survey. In OPL #47, HCFA mandated that all Section 1876 Risk or Cost managed care plans report data to HCFA on 32 selected HEDIS 3.0 measures and participate in the Consumer Assessment of Health Plan Study (CAHPS) survey. These requirements were developed as part of HCFA's initiatives to assist health plans with data collection for use in quality improvement studies, provide Medicare beneficiaries with comparable information on health plans, and assist HCFA in its oversight responsibilities.

As part of its objective to certify data validity, HCFA required that, following submission of the Medicare HEDIS data, some of the required HEDIS measures undergo external validation. To perform the external validation, HCFA released a Request for Proposals to all PROs interested in participating in the validation audit. IPRO, the New York State PRO, was selected as the prime contractor. IPRO further subcontracted with Stratis Health, the Minnesota PRO, and MetaStar, the PRO for Wisconsin, to assist in conducting audit site visits. In addition, The MEDSTAT Group was selected to provide technical assistance and additional staffing to IPRO.

Of the 32 HEDIS 3.0 measures required for reporting to HCFA, five measures were selected for detailed evaluation in the audit. Four Effectiveness of Care measures were selected:

- Breast Cancer Screening,
- Eye Exams for People with Diabetes,
- Beta Blocker Treatment After a Heart Attack, and
- Follow-up After Hospitalization for Mental Illness.

The fifth measure, Frequency of Selected Procedures, was selected from the Use of Services Domain, and included 10 procedures.

Methodology

Audit Design and Data Collection Instruments

The audit methodology was designed by IPRO, in consultation with a Technical Advisory Panel, which consisted of representatives from HCFA, IPRO, NCQA, The MEDSTAT Group, several health plans, and other PROs. The audit consisted of two phases: a desk audit and an onsite audit.

For the desk audit, all health plans that submitted 1996 HEDIS 3.0 Medicare data were mailed a Baseline Assessment to complete. Health plans were required to provide information about their membership, data systems, provider networks, ancillary services (including vendors), and methods and processes used to prepare and submit rates for required Medicare HEDIS measures.

For purposes of accurately reporting HEDIS measures, 15 Medicare contracts that cover large geographic areas were further divided into markets. When the markets were tabulated with contracts, the final number of contract-markets that reported the required measures totaled 284. All 284 contract-markets returned the completed Baseline Assessments.

Health plans that were selected for an onsite audit received a Pre-Onsite Information Request form to complete before the onsite audit. This form required health plans to provide detailed documentation on the information systems and processes they used to prepare their HEDIS reports, and requested copies of flowcharts, computer programs, medical record abstraction instruments, and claims/encounter forms. The Pre-Onsite Information Request provided auditors with additional information to use in preparation for the onsite audit visit.

During the site visit, auditors used an IS Process Audit Interview Guide and detailed Measure Checksheets to conduct the audit and to ensure that all required information was collected. The IS Process Audit Interview Guide contained detailed questions on health plans' information systems and processes. HEDIS measure review was guided by the Checksheets, which contained detailed criteria required for each of the five audited measures.

After the site visits, health plans were given Final Reports, which provided them with narrative descriptions of auditors' findings for the different audit dimensions and final audit designations for each measure.

Sampling Strategy for Selection of Onsite Audits

Because of funding and time limitations, onsite audits were performed for only a sample of contract-markets. Results of the desk audit indicated that for many contract-markets there was insufficient Medicare enrollment to produce accurate rates for the five selected measures. Consequently, the contract-markets for the onsite audit phase were selected through the use of the Probability Proportional to Size (PPS) sampling method rather than through random sampling.

The PPS method weights more heavily those contracts with large enrollments (as determined by the December 1996 Medicare enrollment for each Medicare managed care contract) while still allowing for representation of contracts with smaller enrollments. The selection resulted in the onsite audit of 79 contract-markets. The selected sample represented 65 percent of the December 1996 Medicare managed care beneficiary population.

Onsite Audit Visits

Onsite audits were conducted to follow-up on issues identified in the Baseline Assessment and to more fully determine a health plan's ability to accurately prepare HEDIS measures. Following the 1997 NCQA HEDIS Compliance Audit Standards and Guidelines, the site visits were divided into two sections: an information systems assessment and a HEDIS measure review. The information systems component consisted of interviews and facility walk-throughs with key health plan and vendor staff responsible for claims/encounters, data transfer and entry, membership and provider data, data integration and control, and, when applicable, medical record review (MRR). To review the five selected measures, auditors evaluated the automated and manual processes used to prepare the denominator, the MRR sample (when applicable), and the numerator. A key aspect of the onsite visit was the confirmation of information provided in documents and interviews through live demonstrations of key health plan systems used to support HEDIS.

An onsite audit visit typically lasted from 2 to 5 days. Soon after the completion of a site visit, auditors issued draft audit findings to the health plan for each contract-market audited. Health plans were allowed one week to comment on the report findings. Following the comment period, any of the health plans' suggestions or changes that were deemed appropriate were incorporated into the final reports and then distributed to the health plans.

Defining and Measuring Compliance with HEDIS Specifications

An important goal of the audit was to determine if HEDIS rates were usable in health plan comparison. In this audit, whether or not a particular rate was usable for comparison purposes was determined by how compliant the methods used to prepare the rate were with HEDIS specifications. Compliance was measured from an outcome approach, rather than a process approach, as follows:

- **For Effectiveness of Care Measures:** If the auditors determined that the sum of the errors made by the health plan caused the reported rate to deviate by more than 5 percentage points from the true rate—or if the health plan could not produce information to document the extent of the deviation—then the rate was designated as non-compliant and given an audit designation of “NR”. If the deviation was less than 5 percentage points, then the rate was considered compliant and given an audit designation of “R”.
- **For Frequency of Selected Procedures:** Because this measure is reported in a procedures per thousand format rather than as a percentage, the definition of compliance required a slight modification. If the auditors determined that the sum of the errors made by the health plan caused the number of counted procedures to deviate by more than 5 percent from the true number of procedures—or if the health plan could not produce information to document the extent of the deviation—then the rate was designated as non-compliant. If the deviation was less than 5 percent, then the rate was considered compliant.¹

¹ Note that for the Audit of 1997 Medicare HEDIS 3.0/1998 data, the Report / Not Report threshold for the Frequency of Selected Procedures measure was changed to 10 percent, in compliance with NCQA guidelines.

Auditors used a variety of techniques to determine the amount of bias caused by errors in the processes used to prepare a HEDIS measure. In many instances, health plans were able to simply correct errors that they had made in their computer programming and provide auditors with revised HEDIS rates to assess the amount of bias in the rates reported to HCFA. In other instances, health plans ran queries against their claims/encounter, membership, provider, or ancillary databases to estimate the effect of a particular type of error. Health plans sometimes could document the effect of an error—or a lack thereof—by simply printing out a listing of all the members included in the rate and allowing auditors to manually analyze the listing. Finally, the effect of some errors was documented by ad hoc processes unique to the situation. For example, one health plan estimated the effect of not capturing provider specialty from non-participating providers by using a vendor to compare the providers included in a particular measure to a national directory of provider specialties.

Note that auditors were unable to estimate the effect of a particular error in every case. Errors in denominators were particularly hard to measure, as omitting members from a sample would only bias the rate if the omitted members received the service being measured at a different rate than the included members. Some health plans could not estimate the effect of errors on their reported rate because they did not save copies of the databases that they had used to prepare their HEDIS rates. Other health plans did not or could not make available the resources needed to document the effect of an error, particularly if it meant asking a vendor to conduct extra work or trying to locate a programmer who had recently left the plan. If a health plan could not reasonably estimate the effect of an error, then the measure received a designation of Not Report.

Data Analysis

Data from the Baseline Assessment, Pre-Onsite Information Request, and onsite visits were considered together to determine the extent to which the 1996 Medicare HEDIS data were prepared in compliance with HEDIS specifications and, more importantly, whether those data were usable for health plan comparison.

- Responses to the Baseline Assessment for all 284 contract-markets were entered into a database; all questions were close-ended for ease of data entry and analysis.
- The Pre-Onsite Information Request was completed for all 79 contract-markets that were selected for an onsite audit; all close-ended questions on the Pre-Onsite Information Request were also entered into a database.
- As part of the onsite audit, auditors completed a checklist, which covered 16 key audit dimensions (e.g., age, population, contraindications) reviewed for each audited measure. The specific audit elements, which were verified as part of each dimension, were determined by the different HEDIS criteria for each measure. Auditor responses on the checklist fields were entered into a database and final audit designations were made for each measure.

The most basic goal of the audit analysis was to summarize the audit determinations made for the 79 contract-markets that received an onsite audit. Measure-specific audit designations were assigned using two different scoring systems. The first scoring system, which was designed by NCQA, allows auditors to score measures using one of three audit designations: Report, Not Report, and Not Applicable. The second scoring system was a more detailed system that was developed by IPRO. Tables and graphics that profile the audit findings using each of the two scoring systems—as well as more detailed descriptions of the scoring systems—are included in Chapter 2, Audit Findings—Selected HEDIS Measures.

In addition to profiling the proportion of contract-markets that were in compliance with HEDIS specifications for each of the five audited measures, the analysis focused on characterizing the nature of the problems that led to non-compliance. For example, some problems were determined to be widespread, while others were identified as infrequent and specific to only a small number of health plans. Wherever possible, the effect of information systems and HEDIS preparation problems on HEDIS rates were quantified in terms of magnitude (i.e., how many percentage points of bias were introduced into a rate) and direction (i.e., was the rate biased upward or downward).

Detailed descriptions of problems encountered by health plans when preparing their HEDIS rates are also provided. Following the format of 1997 NCQA HEDIS Compliance Audit Standards and Guidelines, Chapters 2 and 3 of this report profile specific problems with health plan information systems and measure-specific preparation processes.

Auditor Training and Oversight

The onsite audit teams consisted of 2 to 5 individuals with diverse backgrounds, including health care data analysts, programmers, statisticians, systems analysts, and medical record review supervisors. To ensure inter-auditor consistency, all auditors attended a rigorous, four-day training session conducted by IPRO and The MEDSTAT Group. The course detailed the audit process and introduced the data collection instruments that were used during the onsite audit visits.

Inter-auditor reliability among auditors was maintained through supervision of the auditors' site visits by senior IPRO and MEDSTAT staff. The IPRO project manager also attended site visits with several audit teams and provided real-time coaching and feedback. Periodically, IPRO provided all auditors with written updates resulting from auditors' suggestions or discussions with HCFA and NCQA.

The consistency of audit designations and reports was ensured through the assignment of three key IPRO staff to review all final reports and the use of standardized checksheets and scoring tables. The three readers ensured final report consistency among themselves by comparing their assessments of a 6 percent sample of reports. Emphasis was placed on three key areas including tone, content, and organization. Reports with minor problems in these key areas were simply edited and finalized; reports with significant concerns were sent back to the auditors for rewriting and, if necessary, additional follow-up with the health plan. All scoring was reviewed by supervisory staff and checked for logical consistencies against detailed scoring assessments.

Data Validation

Extensive data validation was performed on all information gathered from the Baseline Assessment, the Pre-Onsite Information Request, and Measure Checksheets. Information from all data collection instruments was recorded in databases for use in the final analyses. Numerous automated edit checks were built into the data entry tools to reduce the likelihood of keying errors. In addition, extensive data-cleaning queries were performed, and logical checks were performed to reduce data errors.

Internal quality control (IQC) was performed on all entered data. Depending on the complexity of the database, supervisory staff either reviewed all data entered or reviewed a sizable sample

(minimum 20 percent) for accuracy. Data entry staff maintained at least 98 percent accuracy throughout the data entry phase. In addition, data entry staff worked closely with auditors/analysts to obtain any missing information or resolve discrepancies.

All auditor-completed instruments were validated against Final Reports to ensure the accuracy of all information entered. If any discrepancies were found, auditors were contacted and required to clarify information so that only valid data was entered from all data collection instruments.

In addition to the validations that were performed for data entry, additional validations were done to quantify the accuracy of responses to each question included on the Baseline Assessment. For the 79 contract-markets that received an onsite visit, the audit findings made after the site visit were compared to the information provided in the Baseline Assessment. If discrepancies occurred, the Baseline Assessments were retrospectively modified, where applicable, to reflect auditor's actual findings.

Study Limitations

In reviewing the audit findings, it is important to keep in mind the limitations of the study. These limitations include:

- Only five HEDIS measures, from only two measure domains, were audited and, thus, the audit findings could not be generalized to all HCFA-required measures.
- Because medical records were not reabstracted by auditors, the ability to fully validate the accuracy of health plans' medical record reviews was limited. However, auditors did assess health plan medical record review tools, the medical record sampling and retrieval process, and reviewer training and oversight. Several measures were designated as Not Report based upon a review of the medical record review process alone.
- Onsite audits were performed only on a sample of primarily large health plans and, as a number of reporting problems are plan-specific, the audit findings could not be generalized to all Medicare health plans.
- The accuracy of information documented by providers in submitted claims and encounter forms was not assessed.
- None of the audited measures were given designations of "Not Report" if under-reporting occurred because of incomplete data capture, since this issue was not directly addressed by the 1997 NCQA HEDIS Compliance Audit Standards and Guidelines. Yet, as discussed in Chapter 2, incomplete data capture was found to be a common and significant problem for the sampled health plans, particularly for the Frequency of Selected Procedures measure.
- The audit was conducted after health plans had submitted HEDIS rates to HCFA and NCQA, and consequently, errors identified through the audit could not be corrected.

CHAPTER 2: AUDIT FINDINGS—SELECTED HEDIS MEASURES

Using both NCQA's and IPRO's scoring systems, the audit designations for the Effectiveness of Care measures, as well as the Frequency of Selected Procedures measure, are presented and discussed. Because some of the health plans utilized incomplete data for reporting, the findings on Frequency of Selected Procedures are presented in two formats, one that ignores the effect of incomplete claims/encounter data and one that incorporates incomplete data into the audit designations. This chapter concludes with discussions on the general and measure-specific reporting problems experienced by the health plans.

Effectiveness of Care Measures: Findings Presented Using the NCQA Scoring System

Table 1 presents the frequencies of the audit designations, based on 1997 NCQA HEDIS Compliance Audit Standards and Guidelines, for the four audited Effectiveness of Care measures. An audit designation of "Report" (R) indicates that the measure was fully or substantially compliant with HEDIS specifications. A designation of "Not Report" (NR) indicates that the measure was not in compliance with HEDIS specifications or that the health plan chose not to report a measure, even though reporting was required by HCFA. The audit designation of "Not Applicable" (NA) indicates that the plan did not have sufficient population to report a rate.

Table 1 Audit Designations Using the NCQA Scoring System

| Effectiveness of Care Measures | R | NR | NA |
|--|-------|-------|-------|
| Breast Cancer Screening | 88.6% | 6.3% | 5.1% |
| Beta Blocker Treatment After a Heart Attack | 69.6% | 8.9% | 21.5% |
| Eye Exams for People with Diabetes | 73.4% | 24.1% | 2.5% |
| Follow-up After Hospitalization for Mental Illness | 45.6% | 29.1% | 25.3% |

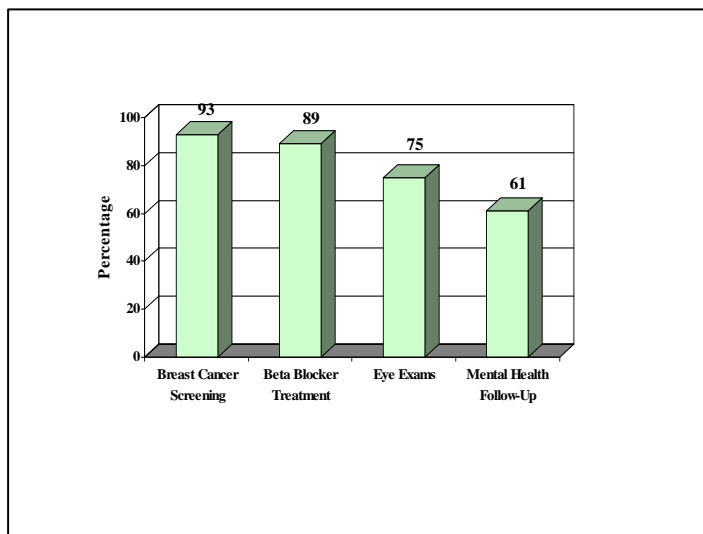
The percentages in Table 1 are based on the 79 contract-markets that were selected for an onsite audit. There were two primary reasons that health plans did not receive a designation of "Report." First, some health plans that were required to report HEDIS rates did not do so accurately or did not report a rate at all and, consequently, received an NR. Second, some health plans did not have sufficient Medicare enrollment to meet minimum sample size requirements and, thus, received an NA.

As indicated by the percentages outlined in Table 1, less than 90 percent of the contract-markets were able to identify an adequate number of enrollees for rate reporting and correctly produce the required rate for any of the four measures. For the Follow-Up After Hospitalization for Mental Illness measure, less than half of the contract-markets were able to produce a reportable rate. Additional analysis of the data used to produce Table 1 found that even though the onsite audit sampling method favored the selection of contracts with large enrollments, some contract-markets did not have sufficient enrollment sizes to produce either the Beta Blocker Treatment

After a Heart Attack or the Follow-Up After Hospitalization for Mental Illness measures, as indicated by the large percentage of NA designations for this measure.

Another way to evaluate the audit findings is to exclude contract-markets with designations of NA and then recalculate the percentage of health plans that accurately produced a rate. Below, Figure 1 illustrates the percentage of contract-markets that reported rates that were fully or substantially compliant with HEDIS specifications, based only on those contract-markets that were required to report a rate for a measure. Specifically, the percentages shown in the graph below were calculated by dividing the number of contract-markets with designations of R by the number of contract-markets with designations of either R or NR, with NA excluded from the calculation.

Figure 1 Percentage of Contract-Markets with Fully or Substantially Compliant HEDIS Data Based Only on Health Plans With Samples Adequate to Report the Measure



Effectiveness of Care Measures: Findings Presented Using the IPRO Scoring System

To more fully characterize the compliance rates of the audited health plans, IPRO developed a detailed scoring system. The audit designations of this system are defined as:

- C** Compliant: Indicates that the plan's systems and processes for preparation of the measure were fully compliant with HEDIS specifications.
- M** Minor Deviation: Indicates that the plan's systems and processes introduced bias into the measure calculation that did not alter the reported rate by more than 5 percentage points.
- S** Significant Deviation: Indicates that the plan's systems and processes introduced significant bias into the measure calculation (i.e., the reported rate was altered by more than 5 percentage points).

NP Not Prepared: Indicates that the plan either did not have in place the systems or processes for reporting the rate or that the plan chose not to report the rate for the measure, even though reporting of the rate was required by HCFA.

NA Not Applicable: Indicates that the plan did not have sufficient population to report a rate.

The techniques used by auditors to determine whether or not a measure was biased are described in Chapter 1, Background and Methodology.

When compared to the NCQA scoring system, the sum of C and M would be equivalent to a designation of R. Similarly, the sum of the IPRO designations of S and NP are equivalent to the NCQA designation of NR. The audit designation of NA was defined by IPRO in the same manner as by NCQA.

Table 2 Audit Designations Using the IPRO Scoring System

| Effectiveness of Care Measures | C | M | S | NP | NA |
|--|-------|-------|-------|------|-------|
| Breast Cancer Screening | 16.5% | 72.2% | 6.3% | - | 5.1% |
| Beta Blocker Treatment After a Heart Attack | 24.1% | 45.6% | 5.1% | 3.8% | 21.5% |
| Eye Exams for People with Diabetes | 32.9% | 40.5% | 21.5% | 2.5% | 2.5% |
| Follow-up After Hospitalization for Mental Illness | 3.8% | 41.8% | 21.5% | 7.6% | 25.3% |

By dividing R into more detailed designation components, C and M, Table 2 demonstrates that the majority of plans with audit designations of R made minor deviations, which were often caused by misinterpretations of HEDIS specifications and information system limitations. Also, by dividing the NR designation into S and NP, it became clear that most plans that received an NR attempted to report a rate for a measure but failed to do so accurately, primarily because of misinterpretations of HEDIS specifications and information system limitations.

Approximately two-thirds of the time, a designation of S was issued because the health plan reported a rate that was more than 5 percentage points above the true rate; the magnitude of over-reporting was generally between 5 to 12 percentage points. About one-third of the S designations occurred because a health plan could not document the magnitude of an error's effect on the reported rate. In only one instance (representing 2 percent of the Not Report designations) did a health plan receive an S designation for reporting a rate that was more than 5 percent below the true rate.

Frequency of Selected Procedures: Findings Presented Using the IPRO Scoring System

Tables 3a and 3b illustrate the audit designations for the Frequency of Selected Procedures measure. These tables used the IPRO scoring system, and because there were no occurrences of NA for this measure, there is no column for NA in the tables.

Unlike the rates for the Effectiveness of Care measures that can be calculated using the hybrid method, which allows for better data capture through medical record review, the Frequency of Selected Procedures measure must be calculated administratively. Consequently, rates for this

measure are more susceptible to underreporting. During the onsite audit visits, it became evident that many health plans used substantially incomplete databases to report this measure.

As the 1997 NCQA HEDIS Compliance Audit Standards and Guidelines did not explicitly specify that the use of incomplete data is sufficient cause to “fail” a measure, this report profiles the audit findings for the Frequency of Selected Procedures in two different formats. Table 3a summarizes the audit designations for each procedure when the use of incomplete data is not considered as a significant deviation. Table 3b shows how the percentages are altered when the true impact of incomplete data is included in determining the audit designations.

**Table 3a: Compliance Scores—
Frequency of Selected Procedures
(Effect of incomplete data not considered)**

| <i>Procedure</i> | C | M | S | NP |
|----------------------------|-----|-----|----|----|
| CABG | 23% | 67% | 9% | 1% |
| Angioplasty (PTCA) | 24% | 72% | 3% | 1% |
| Carotid Endarterectomy | 29% | 68% | 1% | 1% |
| Fracture of Femur | 29% | 66% | 4% | 1% |
| Total Hip Replacement | 29% | 68% | 1% | 1% |
| Total Knee Replacement | 30% | 67% | 1% | 1% |
| Pt. Excision of Lg Intest. | 30% | 61% | 8% | 1% |
| Cholecystectomy, Open | 30% | 66% | 3% | 1% |
| Cholecystectomy, Closed | 30% | 66% | 3% | 1% |
| Hysterectomy | 28% | 68% | 3% | 1% |
| Prostatectomy | 25% | 71% | 3% | 1% |

**Table 3b: Compliance Scores—
Frequency of Selected Procedures
(Effect of incomplete data considered)**

| <i>Procedure</i> | C | M | S | NP |
|----------------------------|-----|-----|-----|----|
| CABG | 23% | 48% | 28% | 1% |
| Angioplasty (PTCA) | 24% | 48% | 27% | 1% |
| Carotid Endarterectomy | 29% | 44% | 25% | 1% |
| Fracture of Femur | 29% | 42% | 28% | 1% |
| Total Hip Replacement | 29% | 44% | 25% | 1% |
| Total Knee Replacement | 30% | 43% | 25% | 1% |
| Pt. Excision of Lg Intest. | 30% | 42% | 27% | 1% |
| Cholecystectomy, Open | 30% | 43% | 25% | 1% |
| Cholecystectomy, Closed | 30% | 43% | 25% | 1% |
| Hysterectomy | 28% | 46% | 25% | 1% |
| Prostatectomy | 25% | 48% | 25% | 1% |

As indicated in Table 3a, contract-markets used measure-preparation processes that significantly deviated from HEDIS specifications less than 10 percent of the time for all eleven of the procedures and less than 5 percent of the time for all but two of the procedures. Yet, when the effect of incomplete data is considered, the scores of significant deviations increase, on average, from 4 percent to 27 percent. For all procedures, when the effect of incomplete data is considered, less than 75 percent of contract-markets reported a fully or substantially compliant rate.

The magnitude of under-reporting because of incomplete data varied tremendously. In many cases, it was clear that incomplete data caused under-reporting of more than 5 percent, but the health plan had little or no documentation for auditors to use to estimate the actual magnitude. When data were available, the effect of incomplete data generally ranged from 10 percent to 40 percent or more.

Problem Areas

The onsite data-collection instrument required auditors to document detailed, measure-specific errors for 16 key audit dimensions. The instrument format profiled measure-specific problem areas by specific audit dimensions for the denominator, sampling and numerator. Denominator dimensions included payer/product, contract/market, age, sex, evidence of service, enrollment, data completeness, and contraindications. For sampling, dimensions included random sampling, unbiased sampling, correct oversampling, and appropriate substitution. The numerator dimensions were time period selected, sufficient evidence of service, data completeness, and, for hybrid methodology, contraindications.

The most common problems experienced by health plans in preparing each of the five audited HEDIS measures are detailed below. It is important to note that the problems listed below sometimes caused only minor deviations in the reported rate but at other times caused significant deviations. Whether a particular problem caused a minor or significant deviation depended upon the systems and processes used by a particular plan, the requirements of the particular HEDIS measure, and the presence or absence of other problems that could have either a reinforcing or mitigating effect. Also, note that the percentages provided in each section below are not mutually exclusive; a particular health plan could have had multiple problems in producing just one measure.

Breast Cancer Screening

The most common problems that health plans encountered in reporting the Breast Cancer Screening measure involved following denominator specifications of continuous enrollment, contraindications, and age identification. Sixty-eight percent of health plans had problems either tracking two years of continuous enrollment or programming for the appropriate denominator population. Thirty-two percent of health plans had difficulty properly applying contraindication criteria. While excluding members having contraindications was optional, once the health plan chose to do so, it had to be done accurately. A few health plans, nearly 12 percent, also encountered problems with the age definitions for this measure. There were no common problems in defining the numerator or in sampling.

Beta Blocker Treatment After a Heart Attack

Most of the problems that health plans encountered with the Beta Blocker Treatment After a Heart Attack measure involved defining the denominator population. Twenty-eight percent of the health plans incorrectly used admission date rather than discharge date. Almost 18 percent of the health plans failed to exclude subsequent AMI episodes of care, while 16 percent incorrectly included members who had discharges that occurred outside of the reporting year. Some health plans failed to use only the first episode of AMI in 1996, or did not ensure that only live discharges were included in the denominator. Twelve percent of the health plans incorrectly applied contraindications to measure criteria for member exclusions.

Eye Exams for People with Diabetes

The problems that health plans experienced with the Eye Exams for People with Diabetes measure were found in the denominator, sampling and numerator audit dimensions. In defining the denominator, nearly one-fifth of the health plans had difficulty applying continuous enrollment criteria, and 12 percent defined a diabetic member using proprietary codes that could not be mapped back to standard codes. For sampling, 10 percent of the health plans incorrectly substituted medical records.

The most common problem with this measure was the application of the provider specialty criterion for numerator inclusion. Nearly half of the health plans did not check that the retinal ophthalmoscopic examination was performed, as defined in the specifications, by an optometrist or ophthalmologist. The failure to apply this criterion resulted in overstatement of the health plans' HEDIS rate for this measure. Similar to the denominator problem, 15 percent of the health plans used proprietary codes that could not be mapped back to standard codes to define the service. Eleven percent of the health plans used incorrect standard codes (e.g., ICD-9, CPT-4) to define the ophthalmic service.

Follow-up After Hospitalization for Mental Illness

Of the audited measures, the Follow-up After Hospitalization for Mental Illness measure was the most difficult measure, as illustrated by the low compliance rates, for the health plans to report. Reporting problems were found in the denominator, sampling and numerator audit dimensions. For the denominator, one-third of health plans used standard codes that were not specified in the HEDIS technical manual. Twenty-nine percent of the health plans did not define the members' age at the time of discharge, which is a required specification. Nearly 20 percent of the health plans used the discharge date instead of the admission date to include members in the denominator, and 18 percent failed to properly address non-mental health/chemical dependency readmissions. Another error that some of the health plans made in defining the denominator population was to include discharges that occurred after the first 330 days of the reporting year.

For sampling, 22 percent of the health plans did not select a non-biased sample as defined by the HEDIS specifications.

Whether using the administrative or hybrid methodology, many health plans encountered numerator problems. For health plans that used the administrative methodology, 42 percent failed to check that the required service (Follow-Up After Hospitalization for Mental Illness visit) was performed by a mental health specialist. Nearly one-third of the health plans that used the hybrid methodology did not follow HEDIS medical record review documentation standards; for example, they allowed providers to attest that a follow-up visit occurred without providing evidence from the medical record. Twenty-eight percent encountered at least some difficulty in obtaining the required mental health/chemical dependency administrative data. Problems in obtaining mental health/chemical dependency data often occurred because of confusion among providers about confidentiality laws and resistance among provider groups, provider trade organizations, or behavioral health vendors. Also, over 25 percent of the health plans used standard codes that were not specified in the HEDIS technical manual.

Frequency of Selected Procedures

In reporting the Frequency of Selected Procedures measure, some health plans experienced significant difficulties. Forty-one percent of the health plans did not calculate member age as of the service or discharge date. Almost one-third of the health plans used standard codes that were not listed in the HEDIS specifications. Twenty-nine percent of the health plans had difficulty following special rules for counting CABG and PTCA services when these services were performed in conjunction with one other (i.e., on the same date). Some health plans incorrectly counted individual arteries for the CABG procedure, or counted the same CABG procedure twice on the same service date. Also, almost one-third of the health plans were unable to link the procedure to the service date.

Largely due to the use of incomplete databases, over one-quarter of the health plans could not report accurate rates for this measure. While most of the process errors described in the previous paragraph led to only minor deviations in the reported rates, incomplete data led to rates that often grossly understated true performance.

For all five of the audited HEDIS measures, the number of errors made when preparing the measure did not correlate directly to whether a measure was deemed compliant. Some health plans made only a single error that led to more than 5 percentage points of bias in a reported rate and, consequently, received an audit designation of “NR”. Conversely, some health plans made many, minor errors with a net effect of 5 percentage points or less bias in a reported rate and received an audit determination of “R”.

CHAPTER 3: AUDIT FINDINGS — INFORMATION SYSTEMS AND PROCESSES

This chapter presents descriptive audit findings regarding health plan information systems and processes. The goal of this chapter is to help HCFA, NCQA, and health plans better understand common information system and process challenges faced by health plans when preparing HEDIS rates.

The findings in this chapter rely on two data sources. As the goal of this analysis is to describe, to the extent possible, the information systems of all health plans that are required to report Medicare HEDIS data, the primary data source for this chapter is the Baseline Assessment, which was completed by all 284 contract-markets. Wherever appropriate, the findings from the Baseline Assessments were confirmed or expanded by using data collected from the 79 contract-markets that received onsite audits. Note that in the sections below we use the term “all contract-markets” whenever we discuss findings based on the Baseline Assessments completed by the entire universe of 284 contract-markets. We use the term “audited contract-markets” to refer to information collected as part of the onsite audit process conducted for 79 contract-markets.

When reviewing the findings in this chapter, it is important to remember that the information systems review conducted as part of this audit was not a general, comprehensive systems review. Rather, it was focused specifically on those information systems and processes that impact a health plan’s ability to report HEDIS.

The most common information systems challenges facing health plans when preparing HEDIS are detailed below. It is important to note that the problems listed below sometimes caused only minor deviations in reported HEDIS rates but at other times caused significant deviations. Whether a particular problem caused a minor or significant deviation depended upon the systems and processes used by a particular plan, the requirements of the particular HEDIS measure, and the presence or absence of other problems that could have either a reinforcing or mitigating effect.

Limitations of Baseline Assessment Data

For the 79 contract-markets that received an onsite visit, the audit findings made after the site visit were compared to the information provided in the Baseline Assessment. If discrepancies occurred, the Baseline Assessments were retrospectively modified, where applicable, to reflect auditor’s actual findings. While this comparative analysis revealed that the accuracy of most Baseline Assessment fields was above 80 percent, information provided in three areas had accuracy rates between 60 and 80 percent. The three areas are knowledge of whether or not (1) vendors are used to provide ancillary services, (2) the health plan uses any internally developed codes, and (3) the health plan already had its 1996 HEDIS data audited by another organization.

In two of these three areas—i.e., whether vendors are used to provide ancillary services and whether internally developed codes are used on claims/encounter forms—the trend was consistently in the direction of understatement. In other words, the Baseline Assessment stated that no vendors and no internally developed codes were used when, in fact, they were. Based

upon feedback from health plan staff during site visits, information regarding the use of vendors and proprietary codes is often compartmentalized within the health plan and, consequently, easily forgotten when completing a survey like the Baseline Assessment. Also, there is much disagreement among health plans about what, exactly, constitutes a vendor.

The omission of information regarding use of vendors and use of proprietary coding is critical, in that problems with vendors and proprietary coding are among the major reasons why HEDIS measures are determined to be inaccurate and, therefore, not reportable. An important consequence of this finding is that detailed audits that include onsite review appear to be needed to ensure that accurate information is obtained regarding key health plan processes that affect HEDIS preparation.

Information Systems and Process Findings

The findings below are profiled using categories that mirror the 1997 NCQA HEDIS Compliance Audit Standards and Guidelines, which include Medical Forms and Coding, Data Transfer and Entry, Medical Record Review, Membership Data, Provider Data, Data Integration and Control Procedures, Adequacy of Documentation, and Outsourced or Delegated HEDIS Reporting Functions.

Medical Forms and Coding

Although nearly all of the 284 contract-markets utilized standard hospital and professional forms, 39 percent of contract-markets also used proprietary forms, including encounter forms. During our onsite visits, we reviewed examples of these proprietary forms and found that several of them did not have space for all digits of diagnosis or procedure codes; in addition, sometimes the provider was contractually required to submit only the first three digits of codes. Moreover, approximately 12 percent of the audited contract-markets did not accurately distinguish between primary and secondary diagnosis codes. Often, the distinction between primary and secondary codes simply reflected the placement of codes on the form (i.e., the first code checked was assumed to be primary), and because codes were sometimes arranged alphabetically on encounter forms, diagnoses and procedures beginning with the first letters of the alphabet sometimes were coded incorrectly as primary.

Fifty-nine percent of all contract-markets used internally developed codes in their systems. For the five HEDIS measures selected for review, use of internal codes varied from 9 percent of all contract-markets that prepared the Beta Blocker Treatment After a Heart Attack measure to 23 percent that prepared the Follow-Up After Hospitalization for Mental Illness measure (accuracy of this finding ranges from 80 to 89 percent). Findings from the onsite audit indicated that several of the mappings to standard codes were performed manually and were poorly documented. The mapping of these codes to standard codes involved either the bundling of services into one all-inclusive code or the use of more detailed coding schemes that had no standard coding equivalent.

Data Transfer and Entry

Many of the 79 contract-markets that received an onsite audit did not receive complete encounter data and most were unable to accurately determine the extent of missing encounter data. Onsite audit findings indicated that poor data capture was due primarily to the lack of appropriate financial incentives in a capitated environment, inadequate information systems, inefficient data collection and processing procedures, and insufficient supervision and training of vendors and service delivery staff.

Of the audited contract-markets, only 19 percent reported that they provided financial incentives for their network providers to submit encounter data. This low rate of contract-markets with financial incentives may be partly the cause of the poor submission rates of encounter data.

Some of the audited contract-markets experienced other problems with their data transfer and entry, including inaccurate processing of claims/encounters, insufficient oversight of claims/encounter vendors, data entry backlogs, and difficulties with system upgrades. In some instances, claims and encounters were processed with key information missing or dummy codes. Many processing systems accepted only a limited number of codes (e.g., one or two secondary diagnosis codes). In these instances, the claims/encounter system was often not programmed to reject claims with missing or incomplete information. Contract-markets that used vendors for ancillary services often conducted little oversight of the vendor's activities and were often uninformed about the vendor's data processing system, edit checks, or processes used to transfer the encounter data to the health plan. Several contract-markets also had sizable claim data entry backlogs that delayed claim and encounter processing.

Approximately 41 percent of the audited contract-markets had recently undergone a system upgrade or consolidation. Not surprisingly, the onsite audit found these contract-markets were more likely to experience difficulties with the collection and transfer of claims/encounter data. In addition, 77 percent of audited contract-markets had contracted with an encounter data intermediary that collects and submits encounter data.

Onsite audit findings also showed that vendors are commonly used for provision of ancillary services including mental health, vision, radiology, laboratory, and pharmacy.

Most of the ancillary service vendors provide member-specific data that are directly integrated into the contract-markets' databases. Health plans' assessments of the completeness and quality of their vendors' data indicated that some had uncertainties about the accuracy of their vendor data. For example, mental health services frequently were outsourced to a vendor, and results of the onsite audit illustrated that mental health vendor data were either incomplete or of poor quality for 36 percent of the audited contract-markets.

Medical Record Review

As discussed in Chapter 1, Background and Methods, the audit did not include a reabstraction of medical records to validate the accuracy of medical record review conducted by health plans. However, auditors did interview health plan staff responsible for supervising medical record reviewers and overseeing the entire medical record review process. The onsite audit included an assessment of each health plan's medical record review tools, the medical record sampling and

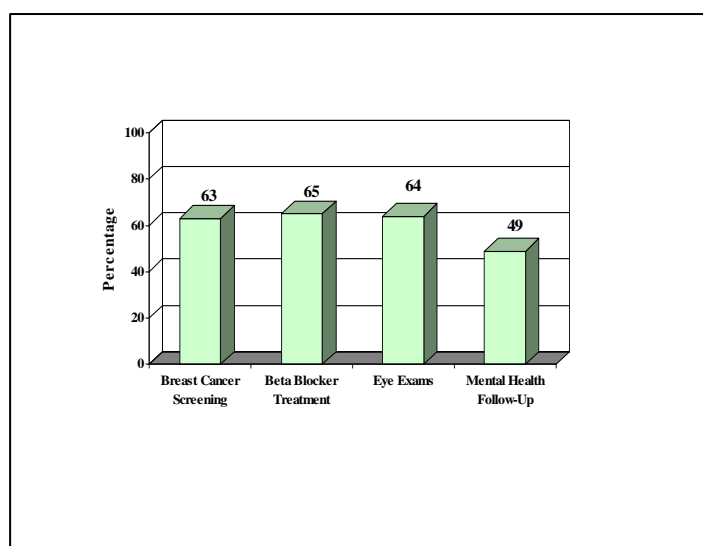
retrieval process, and reviewer training and oversight. Several measures were designated as Not Report based upon a review of the medical record review process alone. Consequently, while a reabstraction of a sample of medical records would have added to the findings in this section, combining information gathered from the Baseline Assessment with findings from the onsite audit of the medical record review process yielded a number of significant findings.

Of all 284 contract-markets, 24 percent did not perform medical record review (MRR) for reporting any of the Medicare HEDIS measures. Of the contract-markets that did perform MRR, nearly 28 percent utilized only their staff employees for the MRR.

During the onsite audit, however, we found that some medical record reviewers received little or no training in HEDIS specifications and that MRR instruments sometimes contained little information on measure-specific criteria to help guide reviewers to abstract the appropriate information. A small number of contract-markets failed to use any standard tools in their MRR; reviewers were simply provided with lists of patient and provider names and required to find evidence of performance of a particular service (e.g., a mammogram). Also, in some instances, MRR information was data entered into a spreadsheet that had no date or code validation checks. Finally, because many of the MRR processes were not documented, there often were insufficient data for inter-reviewer reliability checks or audits.

For three of the audited Effectiveness of Care measures, the hybrid methodology, which included MRR, was used more frequently than the administrative method, as shown in Figure 2, below. Only for the Follow-Up After Hospitalization for Mental Illness measure did the contract-markets use the hybrid method less often than the administrative method, although the difference was slight (49 percent used the hybrid method). Most contract-markets used the hybrid method to overcome administrative system limitations.

Figure 2 Percentage of Contract-Markets using Hybrid Methodology to Report the Measure



Membership Data

While some health plans used Social Security or HIC numbers to identify their Medicare beneficiaries, others used a plan-designated ID number. Importantly, the audit discovered that nearly half of all contract-markets—47 percent—used both forms of identification. Many health plans need to use both forms of identification to link members across multiple systems that were not designed to interact with one another and often identify members using different key identifiers.

During our onsite audits we found that, despite attempts by many contract-markets to upgrade their member enrollment systems, a few contract-markets had significant data entry backlogs in processing of new members. Also, some contract-markets had difficulty tracking members as they moved from one product line to another (e.g., from commercial to Medicare) or when their status changed (e.g., divorce from primary subscriber). Difficulties in tracking members primarily stemmed from how the contract-market identified its members.

For a small number of audited contract-markets, membership data systems did not maintain sufficient lines of history about a member. As additional lines of history were entered, previous history lines were deleted. Plans with these system limitations could not accurately track member histories as required for reporting of several HEDIS measures with continuous enrollment criteria. In other instances, key fields, such as county of residence and gender, were not required to process a membership application.

Provider Data

The onsite audit found that provider data including credentials, tax identification numbers, and locations of service were sometimes maintained on information systems that did not link to the claims/encounter and membership databases. Consequently, much of the linking had to be performed manually. Often, the contract-markets did not sufficiently document the processes by which they performed the links. Furthermore, a few of the provider data systems did not maintain sufficient lines of provider history to accurately produce all required HEDIS measures.

Data Integration and Control

Analysis of the Baseline Assessment found that 60 percent of all contract-markets generated their HEDIS data from programs developed by health plan employees. During our onsite audits, however, we found that many health plans experienced difficulty with staff turnover and adjusting to HEDIS specification changes.

Onsite auditors identified several concerns about appropriate integration and control of data, including insufficient process documentation, inadequate programmer documentation and staffing, system integration problems, and failures to follow standard information system protocols. Some contract-markets, especially smaller ones or those that utilized several vendors that each provided or processed claims/encounter data, often did not maintain sufficient documentation of file integration and system maintenance. Many contract-markets had either insufficient programming staff or had limited documentation of program source code or programming procedures. Several contract-markets were unable to automate system integration of data, especially vendor data, required for HEDIS reporting.

Finally, standard IS protocols were sometimes not followed by information systems staff. Examples of standard IS protocols include version control of files and programs, adequate testing of files and programs, documentation of job logs, and freezing of file extracts.

Adequacy of Documentation

Of the 284 contract-markets that participated in the desk audit, 98 contract-markets or 35 percent had previously undergone an external HEDIS audit. Given that most contract-markets were not familiar with the documentation requirements of a HEDIS audit, it is not surprising that the onsite auditors frequently encountered insufficient documentation.

Auditors found that many contract-markets had limited documentation in many areas, particularly of computer programs, medical record review processes, internal and external reporting, mappings of proprietary codes to standard codes, and processes required for HEDIS reporting.

Outsourced or Delegated HEDIS Reporting Functions

Onsite audits of the contract-markets found that coordination with service and applications vendors was often inadequate. Also, it was often evident that reporting responsibilities were usually not clearly assigned and that oversight of these vendors was insufficient.

CHAPTER 4: CONCLUSIONS

In this chapter, we synthesize the audit findings presented in the previous chapter to draw conclusions that reflect the reality in which HEDIS measures are produced and audited. Conclusions are presented separately regarding the usefulness of the 1996 HEDIS data for health plan comparison, the systems and process limitations that impeded HEDIS reporting, and lessons learned regarding HEDIS audit methods.

Usefulness of 1996 HEDIS Data for Health Plan Comparison

Most health plans were able to produce accurate HEDIS rates for the five measures selected for the Audit of 1996 Medicare HEDIS data. As shown in Figure 1, of health plans with enough enrollees to report a rate, the percentage of contract-markets with reportable rates ranged from a low of 61 percent for the Follow-Up After Hospitalization for Mental Illness measure to a high of 93 percent for Breast Cancer Screening. While for some measures the number of contract-markets that were able to report rates may not be as high as desired by potential users of HEDIS data, most health plans should be recognized as having successfully reported HEDIS rates.

Based upon our findings for a sample of health plans that received an onsite audit, the following conclusions can be made regarding the ability to use particular HEDIS rates for health plan comparison:

The Breast Cancer Screening and Beta Blocker Treatment After a Heart Attack measures were produced in full or substantial compliance with HEDIS specifications by approximately 9 out of 10 contract-markets. Assuming that the accuracy rate is similar for the contract-markets not included in our sample, these measures may be appropriate for use in making health plan comparisons by individuals who are informed of the limitations of the HEDIS data and this audit.

The Eye Exams for People with Diabetes and Follow-Up After Hospitalization for Mental Illness measures were produced in full or substantial compliance with HEDIS specifications by 75 percent and 61 percent, respectively, of the contract-markets required to report these measures. Assuming the accuracy rate is similar for plans not included in our sample, these measures would generally not be considered appropriate for health plan comparison, as there are likely to be many inaccuracies among the HEDIS submissions for health plans with unaudited data.

More than 9 out of 10 contract-markets used processes that are in full or substantial compliance with HEDIS specifications for the Frequency of Selected Procedures measure. Yet, when the effect of incomplete data on the reported rate is considered, only approximately 75 percent of contract-markets prepared a rate that is in full or substantial compliance with HEDIS specifications. Health plans cannot use the hybrid option to overcome the effects of incomplete data for this measure and, consequently, they must report rates that understate true performance if their claims/encounter databases are incomplete. This limitation is particularly important for this measure because, unlike the Effectiveness of Care measures (which are all designed such that a higher rate is better), for the Frequency of Selected Procedures measures, a lower rate might be considered desirable in some situations. Therefore, the consequences of under-reporting could inappropriately favor a health plan when making comparisons to its competitors.

Systems and Process Limitations that Impede HEDIS Reporting

Most health plans—including those that successfully produced a reportable rate—had at least some difficulty reporting the five audited HEDIS measures. Indeed, many of the health plans that ultimately received a designation of Report for a measure had to use a variety of workarounds and manual processes to overcome system and data limitations. Errors that lead to inaccurate HEDIS reporting came from several sources, including limitations of health plan systems, misinterpretation of HEDIS specifications, and random errors in computer programming.

Many health plans currently do not have the information systems that can fully meet the challenges of HEDIS reporting. As described in Chapter 3, Audit Findings—Information Systems and Process, the systems challenges that health plans face are numerous and diverse. Some of the most significant problems include:

- Use of proprietary coding schemes that cannot be accurately mapped to standard coding schemes (e.g., ICD-9-CM, CPT, DRG).
- Limited ability to capture and integrate all required claims and encounter data from providers and vendors, particularly those paid under capitated arrangements.
- Inability to integrate data required for HEDIS reporting from multiple information systems, particularly as a result of patchwork system upgrades or mergers between health plans with incompatible systems.
- Membership data systems that do not support continuous enrollment requirements or prevent health plans from tracking members as they move among payers and product lines within a health plan.
- Provider data systems that lack the ability to accurately supply provider specialty, credentialing information, and other key data needed for HEDIS reporting.

Many health plans will need to invest significant staff time and money to upgrade their systems to make them fully compliant with the needs of HEDIS reporting and, thereby, eliminate the need for the variety of manual processes needed to overcome their system limitations.

Misinterpretations of HEDIS specifications still cause some health plans to produce HEDIS rates that, when audited, are determined to be non-compliant with HEDIS specifications. While at some health plans all staff involved in the HEDIS preparation process meet regularly to ensure that they have an accurate and consistent understanding of HEDIS, at other health plans the process is less coordinated, which often results in miscommunications or misunderstandings regarding HEDIS specifications. In some instances, trends in misinterpretations have been identified across health plans. These trends identify areas where NCQA can work to improve the accuracy and clarity of HEDIS specifications to reduce the possibility of misinterpretation in future years.

A final important source of error in the preparation of HEDIS rates are random errors that go unnoticed because, at many health plans, the amount of time spent on supervising and documenting the HEDIS preparation process is inadequate. The lack of supervisory review means that mistakes often go undetected until an auditor discovers them, at which time it may be too late to correct the mistakes before the HEDIS submission deadline. Inadequate

documentation not only makes it difficult for supervisors and auditors to review the accuracy of a health plan's HEDIS preparation processes, but it diminishes the ability of a health plan to document and learn from its mistakes and improve upon its HEDIS preparation processes each year. For example, several health plans with poor documentation have had to redo large portions of their HEDIS computer programming after the sudden departure of a key programmer for whom there was no back-up programmer who could continue the departing programmer's work.

Lessons Learned Regarding HEDIS Audits Methods

The processes used in this audit can be improved in several ways. One key limitation of this year's audit was the absence of a medical record validation component, which limited the ability of the audit to fully validate whether a measure prepared using the hybrid method was compliant with HEDIS specifications. A validation of a sample of records for at least some of the measures prepared using the hybrid method would provide auditors with important information to supplement their audit findings regarding the medical record review process.

Because data reported on the Baseline Assessment proved to be inaccurate for several key aspects of HEDIS reporting, and because many issues that affect a health plan's ability to produce HEDIS are unique and complex, onsite auditing appears to be required to ensure the accuracy of HEDIS rates. Had every Medicare contract-market been part of the onsite audit process, HCFA would have been able to release rates even for those measures where only 60 or 70 percent of contract-markets received a designation of Report, as the audit designation for every contract-market would be known. This year's sampling approach precluded the release of any information for measures with poor compliance rates because, for the approximately 70 percent of contract-markets that were not audited, the contract-markets with compliant rates cannot be identified and separated from those with non-compliant rates.

Although this year's audit provided HCFA, NCQA, and health plans with a substantial amount of information regarding the problems that health plans encounter when preparing HEDIS rates, the information came too late for health plans to attempt to correct the errors that they made. As described in Chapter 2, some of the errors identified by auditors were easily correctable. Consequently, a pre-submission HEDIS audit would serve not only to identify errors in HEDIS reporting but also to facilitate the correction of many of those errors before rates were submitted to HCFA.

While not a formal part of the audit standards for this Medicare HEDIS audit, incomplete claims and encounter data is clearly an issue that affects the comparability of HEDIS rates, particularly for measures where the hybrid option is unavailable. Clearer guidelines on how to address incomplete data in future audits are required.

Similarly, the accuracy of provider coding on claims and encounter forms is currently beyond the scope of a HEDIS audit. The potential effect of any inaccuracies caused by inappropriate coding of claims and encounters, therefore, remains unknown.

CHAPTER 5: RECOMMENDATIONS

This final chapter presents IPRO's suggestions on the appropriate use of HEDIS data, including recommendations on public release of measure-specific results. Recommendations are also made to Medicare health plans, NCQA, and HCFA regarding process changes that could help improve the accuracy, completeness, and timeliness of Medicare HEDIS reporting. The chapter concludes with suggestions for increasing the effectiveness and efficiency of future Medicare HEDIS audits.

Recommendations for Release of Measure-Specific HEDIS Data

Our recommendations regarding the release of data for the five Medicare HEDIS measures audited by IPRO varies by measure, as follows:

Breast Cancer Screening and Beta Blocker Treatment After A Heart Attack

These measures were found to be prepared in full or substantial compliance with HEDIS specifications by 93 and 89 percent, respectively, of the audited contract-markets that reported a rate. If HCFA deems this level of accuracy to be adequate and, therefore, chooses to publicly release comparisons of health plan performance on these two indicators, IPRO recommends that a clear statement of data limitations be included. Such a statement is particularly important because 205 of the 284 contract-markets did not undergo an onsite audit and, consequently, some biased rates would be reported. In addition, the accuracy of medical record review was not validated for the approximately 60 percent of health plans that used the hybrid method to prepare these measures.

Eye Exams for People with Diabetes and Follow-Up After Hospitalization for Mental Illness

Less than 80 percent of the contract-markets in the audit sample prepared the Eye Exams for People with Diabetes and Follow-up After Hospitalization for Mental Illness measures in full or substantial compliance with HEDIS specifications. Due to the high rate of failure and the numerous difficulties experienced by some health plans in reporting these two measures, IPRO does not recommend that HCFA use these measures for health plan comparison.

Frequency of Selected Procedures

IPRO also recommends that HCFA not use the Frequency of Selected Procedures measure for health plan comparison. While the processes used to prepare this measure were found to be compliant with HEDIS specifications for over 90 percent of the contract-markets in the audit sample, rates for approximately 23 percent of the contract-markets were prepared using databases that were substantially incomplete. Consequently, many health plans' reported rates understate true performance.

Unaudited Measures

Based upon our findings for the five audited measures—as well as our findings regarding health plan information systems and HEDIS preparation processes—IPRO recommends that HCFA be extremely cautious in using any unaudited data for any type of health plan comparison.

Recommendations for Improving Medicare HEDIS Reporting

Medicare Health Plans

To improve the accuracy, completeness, and timeliness of Medicare HEDIS reporting, we recommend that health plans, as well as HCFA, carefully review IPRO's audit findings regarding both general limitations of health plan information systems and measure-specific issues that result in reporting problems. Based upon our audit findings, we recommend particular attention to the issues outlined below:

- Claim and encounter forms, as well as the corresponding databases into which the forms are entered, should capture all information required for HEDIS reporting, particularly regarding diagnosis and procedure codes.
- All health plan forms and databases should capture and validate diagnosis and procedure codes at their highest level of specificity and properly distinguish primary and secondary codes.
- Health plans that use proprietary codes should verify the accuracy of their mapping to standard codes and, where necessary, consider revising their coding practices to be consistent with standard codes used in HEDIS.
- Health plans should study the completeness of their claims and encounter databases and take actions, as necessary, to ensure that HEDIS data accurately and fully reflect health plan performance, particularly for measures for which the Hybrid method is not an option.
- Oversight of data vendors should include ensuring that vendor data quality standards and processes are (1) adequate to support HEDIS reporting and (2) enforced by the vendor.
- Medical record reviewers, whether they are employees of the health plan or contracted from a vendor, should be provided with adequate training, ongoing oversight, and data collection tools that guide them to collecting appropriate and accurate information.
- Membership data systems should be able to support all HEDIS requirements for calculating continuous enrollment, identifying eligibility for specific benefits, and tracking enrollees even as their payers and providers change.
- Provider data systems should adequately track key provider information required for HEDIS, particularly regarding provider specialty and board certification status. In addition, health plans should study their ability to identify provider specialty for claims received from non-participating providers and the effect on HEDIS reporting.
- When preparing their HEDIS measures, health plans should have a detailed and well-supervised process for ensuring that each detail of the specifications for every HEDIS measure is properly implemented.

- Health plans should review Chapter 2 of this audit report to identify errors typically made by health plans when preparing HEDIS measures and then ensure that their HEDIS preparation processes are free of these errors.

Throughout the year, health plans should maintain ongoing communication with staff, providers and vendors about the importance of HEDIS reporting. Some of these groups may not be familiar with HEDIS specifications or the requirements for accurate reporting. Health plans should educate and oversee staff, providers and vendors throughout the year to ensure that data are collected, managed, and transmitted in a manner consistent with HEDIS specifications.

The importance of keeping clear and accurate documentation of all HEDIS activities cannot be overemphasized. Health plans should produce an audit trail that accurately and thoroughly describes the information systems and manual procedures used to prepare their HEDIS reports. To create the audit trail, all process steps should be documented and updated in a systematic and controlled manner. Standardized record keeping of all events relevant to HEDIS reporting should be encouraged.

National Committee for Quality Assurance

Based on our audit findings and feedback from health plans and field experts, there are several recommendations for NCQA to consider to help improve the quality of reported Medicare HEDIS rates. These include:

- During its revisions of specifications, NCQA should consider the impact that the common data limitations, programming difficulties, and measure-specific misinterpretations identified in this audit report may have on the accuracy of the reported HEDIS rates.
- The HEDIS specifications should be written with greater detail. For example, all required codes for the measures should be listed individually and not as ranges (i.e., as 99201, 99202, 99203, 99204, 99205 and not as 99201 – 99205). More detailed specifications should reduce the chance that health plans misinterpret the specifications or make programming errors.
- To help health plans produce quality HEDIS reports, NCQA should consider earlier release of the HEDIS technical specifications manual and HEDIS reporting software.

Health Care Financing Administration

HCFA should consider the following suggestions to maintain its goal of obtaining accurate, comparable, and timely HEDIS data from Medicare health plans:

- Continue to protect the confidentiality of proprietary information obtained from health plans. Health plan staff were more likely to be cooperative when they understood that proprietary information, which may give the health plan a competitive advantage, would not be disclosed to the public.
- HCFA should continue to provide technical assistance to health plans. The assistance should focus on continuous improvement of the quality and completeness of HEDIS data reported to HCFA. At a minimum, the findings of this audit report should be disseminated as quickly and broadly as possible to encourage improved reporting.

- HCFA should continue to be an active participant in the development, implementation, and oversight of policies that improve the accountability of Medicare managed care plans through the use of health care quality measurement mechanisms like HEDIS. As the largest single payer of health care, HCFA has the unique opportunity to lead the effort to identify and address key issues regarding the accuracy and comparability of performance data across health care organizations.

Recommendations for Future Audits

To provide HCFA and Medicare beneficiaries with even greater confidence in the accuracy and comparability of HEDIS data, HCFA may want to consider the following in determining its requirements for future audits:

Because the information systems and HEDIS preparation processes used by each health plan are unique, future HEDIS validation efforts should require an audit of every HCFA contract-market. An audit of this magnitude would allow HCFA to make meaningful comparisons across health plans and gain a better understanding of overall health plan performance.

Because the audit of the selected 1996 HEDIS Medicare measures did not include provisions for health plans to correct their rates, HCFA's ability to present accurate HEDIS rates to Medicare beneficiaries was limited. Future audits should be prospective so health plans have an opportunity to correct their HEDIS rates before submission.

While the audit of 1996 Medicare HEDIS examined the processes health plans used to conduct medical record review, it did not include reabstraction of medical records to verify HEDIS rates for measures prepared using the hybrid method. Because most health plans used the hybrid method for at least one—and usually several—HEDIS measures, the incorporation of medical record reabstraction into future audits would provide HCFA with even greater confidence in the accuracy of reported HEDIS rates.

Because findings from the audit of 1996 data indicated that incomplete encounter databases were a common problem, HCFA may want to require greater attention to the effect of this issue on HEDIS rates in future audits of Medicare HEDIS[®] data.

A pilot study on provider coding practices could provide useful information regarding whether the diagnosis and procedure codes reported via claims and encounter data accurately reflect the information recorded by providers in a patient's medical record.