Use of Performance Information for Quality Improvement

The Role of Performance Measures for Improving Quality in Managed Care Organizations

Dennis P. Scanlon, Charles Darby, Elizabeth Rolph, and Hilary E. Doty

Objectives. To understand how managed care plans use performance measures for quality improvement and to identify the strengths and weaknesses of currently used standardized performance measures such as the Health Plan Employer Data and Information Set (HEDIS®) and the Consumer Assessment of Health Plans (CAHPS®) survey.

Data Sources/Study Setting. Representatives (chief executive officers, medical directors, and quality-improvement directors) from 24 health plans in four states were surveyed. The overall response rate was 58.3 percent, with a mean of 1.8 respondents per plan.

Study Design. This exploratory qualitative research used a purposive sample of respondents. Two study authors conducted separate one-hour tape-recorded telephone interviews with multiple respondents from each health plan.

Principal Findings. All managed care organizations interviewed use performance measures for quality improvement but the degree and sophistication of use varies. Many of our respondent plans use performance measures to target quality-improvement initiatives, evaluate current performance, establish goals for quality improvement, identify the root cause of problems, and monitor performance.

Conclusion. Performance measures are used for quality improvement in addition to informing external constituents, but additional research is needed to understand how the benefits of measurement can be maximized.

Key Words. CAHPS, HEDIS, managed care, NCQA, performance measurement, quality improvement, report cards

As enrollment in managed care health insurance plans has grown in the United States over the past decade for commercial, Medicaid, and Medicare

populations (KPMG 1998; Kaiser Family Foundation 1999; HCFA 1997, 1999), so have concerns about the access, service, value, and quality of care provided by these plans (Peterson 1999). Stakeholders have responded by demanding accountability from managed care plans, often in the form of measurable, standardized performance measures. Performance measures are broadly defined to include quantifiable and objective measurements that can be used to evaluate some aspect of the services provided by managed care organizations (MCO) (Coltin and Aronow 1995).

External stakeholders may use performance measures for various reasons. Some public and private purchasers have developed and supported measurement systems such as the Health Plan Employer Data and Information Set (HEDIS®) (NCQA 1995), the Consumer Assessment of Health Plans (CAHPS®) survey (Carman, Short, Farley, et al. 1999), and plan or network accreditation (Scanlon and Hendrix 1998) to make informed purchasing decisions on behalf of their beneficiaries. Consumers may use performance measures, such as health plan "report cards" routinely published by the media or consumer organizations, to determine in which managed care plan they should enroll. Regulators, such as state insurance or health departments, may use performance measures to ensure that minimal standards of acceptability are met. Ironically, the current literature suggests that these external stakeholders may use performance measures to drive quality improvement much less often than has been predicted by proponents (Gabel, Hunt, and Hurst 1998; Chernew and Scanlon 1998; Scanlon and Chernew 1999).

Other potential users of performance measures are the MCOs themselves. Although many of the performance-measurement systems were developed to inform external constituents, MCOs may find the measures useful for improving the quality of clinical care and service they provide. It is also possible, however, that the measures demanded by external constituents do

---

This project was supported through a subcontract from RAND with funds provided by a grant from the Agency for Healthcare Research and Quality (AHRQ) (grant no. HS09204). The views expressed in this paper are those of the authors and do not necessarily reflect the opinions or position of AHRQ or the institutions with which the authors are affiliated.

Address correspondence to Dennis P. Scanlon, Ph.D., The Pennsylvania State University, Department of Health Policy and Administration, and Center for Health Policy Research, 119B Henderson Building, University Park, PA 16802. Charles Darby, M.A. is from the Agency for Healthcare Research and Quality, Rockville, MD. Elizabeth Rolph, M.A. is from RAND, Santa Monica, CA. Hilary E. Doty, M.S. is from the Department of Health Policy and Administration, The Pennsylvania State University. This article, submitted to *Health Services Research* on April 24, 2000, was revised and accepted for publication on August 31, 2000.

not provide the type of information required and that MCOs will develop their own measures to make meaningful improvements in quality. In this article we seek to determine whether MCOs use performance measures for quality improvement, and if so, what role performance measures play in quality-improvement programs. Our study is exploratory, so our aim is to develop questions for future research rather than to generate generalizable findings for current policy debate.

LITERATURE REVIEW

Few studies have broadly examined how MCOs incorporate performance measures into their quality-improvement programs. Most of the existing studies discuss specific quality-improvement initiatives, focusing primarily on impetus and design rather than the role of measurement (Goverman 1994; Kinney and Gift 1997). Published literature that systematically examines the role and value of HEDIS and CAHPS for purposes of quality improvement is also scarce. One useful source is a compendium of quality improvement "best practices" produced by the National Committee for Quality Assurance (NCQA) and Pfizer Corporation (NCQA 1999a), which highlights 38 quality-improvement initiatives selected by a panel of experts from information gathered by NCQA as part of its accreditation program. It reports on the motivation, design, implementation, and outcome of each intervention. Although the role of performance measures, particularly for HEDIS, is discussed, the report primarily focuses on the design and implementation of the quality-improvement initiatives rather than the role of performance measures for these initiatives.

A study by Landon and Epstein (1999) examined the quality-management practices of a representative sample of Medicaid managed care plans. Although these authors focused on the organizational characteristics of these plans, they did ask respondents if their plans collected and used six performance measures, such as childhood immunization rates and first trimester prenatal care, that are relevant to the Medicaid population. Their findings revealed that most plans collect these measures and that usage does not vary significantly for Medicaid and commercial plans; however, their study was not designed to provide details on the manner and extent to which performance measures are used for quality improvement. In another study by Landon, Tobias, and Epstein (1998), representatives from state Medicaid agencies were asked whether plans were required to target specific

performance measures for quality improvement, and if so, whether they had demonstrated significant improvements in these areas. The results suggest that plans fare much better at targeting than improving; once again, however, the study was not designed to explore in detail how performance measures are used for quality improvement.

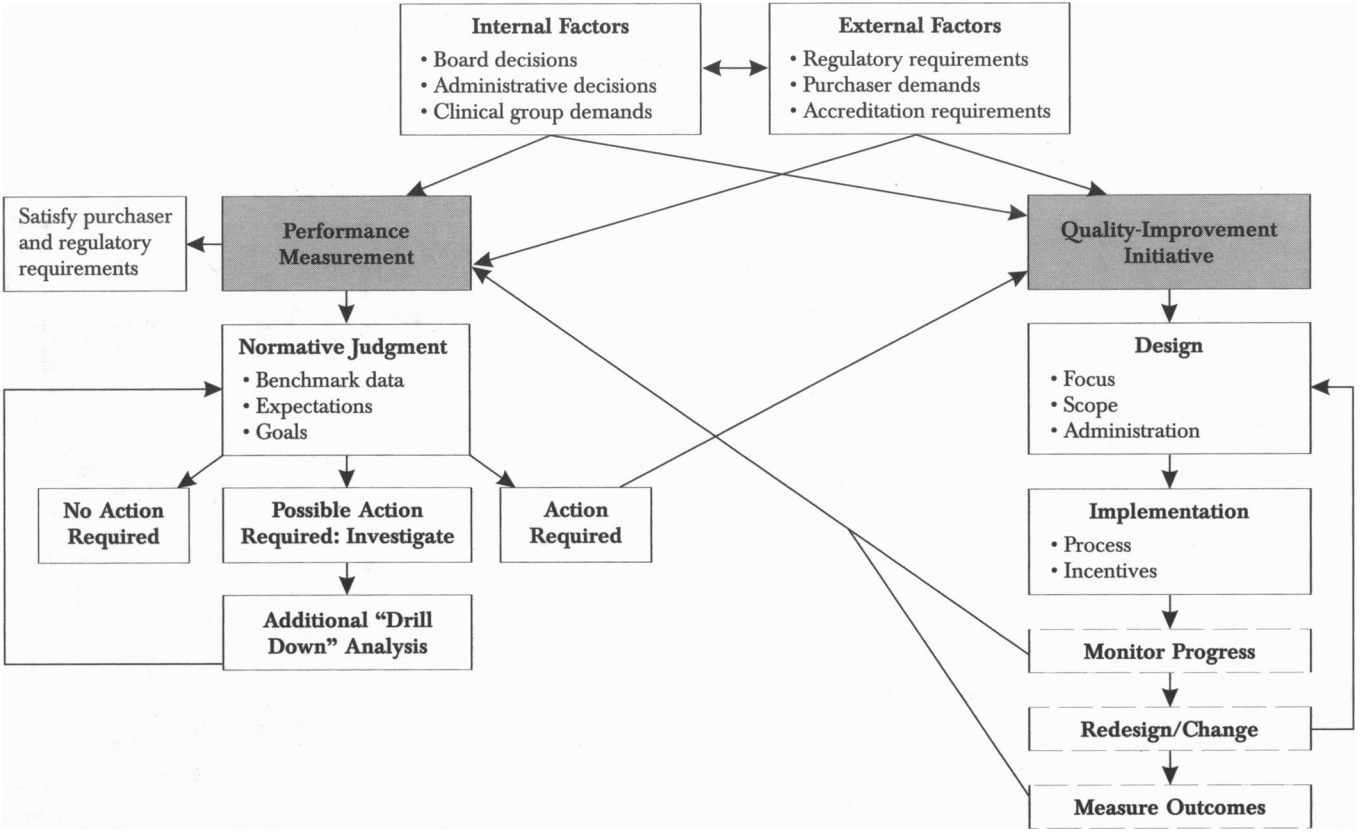
Finally, a study by Hillman and Goldfarb (1995) examined the quality-improvement programs of six HMOs selected from a list of 31 HMOs deemed exemplary by a panel of experts. The authors conducted site visits in 1992 to identify characteristics that made these quality-improvement programs exemplary. Although their study predated the widespread adoption of HEDIS and other standardized performance measurement systems, the authors found that the six organizations were significantly ahead of other plans in terms of collecting data that are useful for quality-improvement programs.

CONCEPTUAL FRAMEWORK AND RESEARCH QUESTIONS

Using the existing literature as our guide we developed a conceptual framework for relating quality improvement and performance measurement in MCOs. As Figure 1 illustrates, the decision to engage in performance measurement or to conduct a particular quality-improvement initiative can be driven by internal or external factors or both. The right side of Figure 1 represents the basic steps in any quality-improvement initiative (Jaeger, Kaluzny, and McLaughlin 1994). We hypothesized that performance measures would be useful for monitoring the progress of particular quality initiatives and for measuring the outcomes of initiatives.

The left side of Figure 1 represents our conceptualization of how an MCO might use performance measures. As indicated, an MCO may simply engage in measurement to satisfy purchaser or regulatory reporting requirements, although such requirements increasingly include the stipulation that quality-improvement activities be documented. We hypothesized that once an organization has measured some aspect of its performance, decisions regarding the need for further action may be based on internal expectations or external norms such as comparison with other plans reporting HEDIS data in NCQA's *Quality Compass* (NCQA 2000a). In some cases the available measures may identify poor performance but may not identify the reason for the substandard performance, or the measures may not provide enough information to facilitate the design of a quality-improvement initiative. In

Figure 1: Linking Measurement and Improvement



these cases further analyses will be required before a decision can be made about whether an initiative should be developed.

Because very little research has documented the use of performance measurement for quality improvement by MCOs, our study was guided by this framework, was exploratory in nature, and was designed to address the following research questions.

- Do MCOs use performance measures to improve quality?
- What role do performance measures play in quality-improvement initiatives?
- What characteristics of performance measures make them more or less useful for quality improvement?
- Because many purchasers and regulators specifically require plans to report HEDIS and CAHPS data, we were also interested in the following question: For assisting with quality-improvement activities, what are the strengths and weaknesses of HEDIS and CAHPS as examples of standardized performance measures?

STUDY DESIGN AND METHODS

Because the use of performance measures by MCOs is an unexplored topic, we chose a qualitative, exploratory research design that used interviews with executives in a small, purposive sample of managed care plans. Hence, the findings from our study are not generalizable to health plans in the United States or to the plans in the states where we conducted our study. However, this design allowed us to explore in depth some relatively new areas with representatives of health plans who had considerable experience in the use of performance measurement and its application to quality improvement. Quality improvement can be broadly defined as the systematic and ongoing effort to provide the most effective clinical care and customer service in a manner that best meets the needs of patients and plan members (McLaughlin and Kaluzny 1994).

Sample

Because this was an exploratory study we selected plans that had been exposed to some form of public reporting of their performance scores to increase the likelihood that they would be aware of performance measurement and the standardized tools, particularly HEDIS and CAHPS. Because of budget limitations and because we suspected that geography and state regulation might well influence the use of performance measures, we selected several

plans in each of four states (Pennsylvania, Maryland, Kansas, and Washington) where there had been statewide public reporting of plan performance with HEDIS and CAHPS. Choosing states with exposure to CAHPS was an important criterion for site selection. The Agency for Healthcare Research and Quality, which funded CAHPS and our research, is interested in whether and how CAHPS is used for quality-improvement purposes.[1] Within the four states we chose from larger health plans, plans that were diverse with respect to model type, and plans that served Medicaid, Medicare risk, public employee, and commercial populations.

Choosing states engaged in performance reporting and choosing larger plans within those states reflects our desire to interview plan representatives who were likely to have a concentration of experience not found in a sample of plans from which generalizable conclusions could be made. However, selecting plans that had experience with HEDIS and CAHPS is not necessarily limiting because their use is fairly widespread and not constrained to states with public performance reports. Although we make no attempt to generalize beyond our study population, Table 1 compares the rates of commercial, Medicare, and Medicaid managed care penetration in the study states with national data.

Protocol Development and Participation

We sought information from the chief executive officer (CEO), medical director, and director of quality improvement from each MCO in our sample. We expected medical directors and directors of quality improvement to

Table 1: Managed Care Penetration

State	*Total Managed Care Penetration (%)**	*Medicare Managed Care Penetration (%)†*	*Medicaid Managed Care Penetration (%)‡*
Kansas	16.8	7.4	49.6
Maryland	46.0	11.6	69.6
Pennsylvania	33.6	27.8	70.2
Washington	17.3	24.4	99.9
National range	0–52.9	.2–39.6	0§–100.0
National mean	30.1	17.2	55.6

**Source:* Interstudy (1999).

†*Source:* HCFA (2000a).

‡*Source:* HCFA (2000b).

§Alaska, Virgin Islands, and Wyoming have zero Medicaid managed care penetration.

have operational information and the CEOs to shed light on overarching corporate considerations and strategies. After developing and pilot testing a set of interview protocols tailored to each type of respondent (with four plans from New Jersey) we integrated them into a single instrument that could be administered to all respondents. Use of a single instrument permitted us to corroborate information across respondents and compare perspectives across functions.

Initially we contacted the CEOs, directors of quality improvement, and medical directors of six plans in each of the four states. Five plans, three in Pennsylvania and two in Maryland, refused to participate. Additional plans were recruited to bring the total number of participating plans to 24. We retained a plan in our sample if we were able to secure an interview with at least one of the three principals. There were no noticeable differences between plans that refused to participate and those that agreed to participate other than the fact that no plans in Washington or Kansas declined participation. In all, we interviewed eight CEOs (response rate 33.3 percent), 19 directors of quality improvement (response rate 79.2 percent), and 15 medical directors (response rate 62.5 percent), with an overall response rate of 58.3 percent. The mean number of principals interviewed per plan was 1.8. Of the MCOs we interviewed, 24 (100 percent) offered an HMO plan, 16 (67 percent) offered a point-of-service plan, ten (42 percent) offered a preferred provider organization plan, and seven (29 percent) offered an indemnity plan. Table 2 provides information about the enrollment, model type, populations served, and national affiliation for HMOs that employed our respondents.

Protocol Content

The final interview protocol covered a number of topics including confirmation of our information about MCO characteristics and questions about the respondent's job responsibilities and tenure. We then asked the respondents to describe how quality-improvement activities were organized including details about planning and personnel. Next, we asked respondents to describe three specific quality-improvement initiatives including what prompted the initiatives, whether specific goals have been established, how the progress of the initiatives is being monitored, barriers the initiatives are facing, what measurement tools are being used, and the strengths and weaknesses of these measurement tools. We closed the interviews by asking respondents to evaluate HEDIS and CAHPS and by asking for comments, suggestions, and criticisms for policymakers, regulators, and those who create performance-measurement systems. The interview protocol was standardized, but inter-

Table 2: HMO Plan Enrollment, National Affiliation, and Model Type

*Population**	*Mean Enrollment*†	*Minimum Enrollment*	*Maximum Enrollment*	*National Affiliation (%)*‡
Commercial ($N = 21$)	201,394	7,095	851,849	52
Medicare ($N = 21$)	33,306	1,597	130,251	57
Medicaid ($N = 13$)	35,647	3,089	82,658	46
Total ($N = 24$)	240,769	18,978	969,696	50

Plan Model Type§	*Frequency*	*Proportion (%)*
Staff	0	0
Group	4	17
Independent practice association	12	50
Network	2	8
Mixed	6	25

**Source:* InterStudy (1999); interview data; and NCQA (2000b).

†*Source:* InterStudy (1999). Respondent interview data were used for two plans because InterStudy data were unavailable.

‡*Source:* interview responses.

§*Source:* interview responses. InterStudy data were used to determine the model type for three plans: two plans had within-respondent disagreement and one plan had missing data. Where InterStudy data were unavailable for one plan, the respondent majority was used. One other plan with disagreement was coded as a group model because this plan had only recently expanded and was historically a group model HMO.

viewers probed respondents' answers to obtain more detail and to corroborate responses across respondents from the same plan. Each interview took approximately one hour.

The fact that we interviewed multiple respondents from the same plan helped to increase the number of initiatives we learned about for each organization. Although we asked respondents to choose initiatives about which they could provide rich detail regarding the use of performance measurement, the initiatives chosen may not be representative of all quality-improvement initiatives within each organization. However, we did ask for examples of less successful initiatives so respondents would not limit their choices to successful activities.

Analysis

Two members of the research team conducted each interview in a conference telephone call with the respondent. Respondents from the same health plan were interviewed separately. The interviews were tape recorded, and one interviewer drafted notes from the tape recordings and gave these notes to the

other interviewer to review for accuracy. The final version was used to create a detailed spreadsheet entry for each interview. The spreadsheet facilitated frequency counts and calculations for quantifiable data and aided in sorting and grouping interviews for qualitative analysis. The reported findings were obtained by consensus after several discussions among study authors.

FINDINGS

A total of 42 interviews were conducted, resulting in discussion of 116 quality-improvement initiatives.[2] Table 3 groups the quality-improvement initiatives we discussed with our respondents by category. The majority of initiatives we heard about were in the general areas of preventive care, chronic illness/disease management, and customer service/member satisfaction. Within these categories the most frequently mentioned programs pertained to immunizations, cancer screenings, and care for pregnant women (80 percent of preventive care initiatives mentioned). Diabetes, asthma, and cardiovascular disease were mentioned most frequently for chronic illness (72 percent); enrollee satisfaction, telephone abandonment, member communication, and claims processing were the most frequently discussed initiatives for customer service/satisfaction (76 percent).

Prevalence of Performance-Measurement Use

All MCOs interviewed use performance measures for quality improvement, but the degree and sophistication of use varied. We would have liked to identify characteristics of plans that made exceptional use of performance measures for quality improvement, but our exploratory research design prohibited us from generalizing beyond our study sample. Still, our findings can be used to frame questions for future research.

Affiliation with a large national MCO was a characteristic common to several plans deemed significant users of performance measures.[3] For these plans it was not uncommon for the national office to have established specific reporting criteria and goals. In addition the national organization often provided technical expertise and consultation in the form of data reporting, benchmarking, and quality program development. In many cases this expertise resulted in well-established information systems and data collection processes (Scanlon et al. 2000).

Several of the group model plans appeared to have sophisticated quality-improvement programs that heavily used performance measures. However,

Table 3: Quality-Improvement Initiatives

Category of Quality-Improvement Initiative (N = *116*)	
Preventive care	31%
Chronic illness/disease management	34%
Customer service	25%
Member satisfaction	10%
Specific Preventive Care Initiatives (N = *36*)	
Immunization	33%
Cancer screenings	28%
High risk pregnancy/prenatal care	19%
Other prevention	20%
Specific Chronic Illness/Disease Management Initiatives (N = *39*)	
Diabetes	36%
Asthma	21%
Cardiovascular disease	15%
Mental health	13%
Other chronic illness/disease management	15%
Specific Customer Service/Member Satisfaction Initiatives (N = *41*)	
Improving customer service/member satisfaction	35%
Telephone abandonment rates	17%
Plan-member correspondence/communication	12%
Claims processing	12%
Enrollment card distribution	7%
Provider-member communication	5%
Referral process improvement	5%
Member access to primary care physicians	5%
Reporting lab results by telephone	2%

it is difficult for us to determine what it is about these organizations that caused them to link measurement and improvement because there are many confounding factors. For example, they have more closely linked relationships with their providers, which may enhance cooperation and collaboration between the plan and its physicians. In addition many of these plans are affiliated with a national MCO and have been in existence longer, so their sophistication may simply be because of experience. Finally, all of these plans are not-for-profit plans, suggesting that profit status may be a characteristic of plans that are successful users of performance measures. Disentangling the effect of these important correlates is a subject for future research.

Role of Performance Measurement for Quality Improvement

Our analysis suggests that performance measures are used for targeting quality-improvement initiatives, evaluating current performance, goal setting,

identifying the root cause of problems, and monitoring initiative progress. Although these five categories represent ways that performance measures are used, not all are used in each of these ways. Moreover, plans may use both externally and internally developed performance measures for each of these purposes.

Targeting. Respondents provided many reasons for embarking on the quality-improvement initiatives described. Performance measures were identified as a precipitating factor in 77 percent of the cases and appear to be used in two ways to target quality-improvement activities.[4] First, plans may decide to develop quality-improvement initiatives in areas where measurement is the focus of purchasers, such as Medicaid programs, or accrediting bodies, such as NCQA. However, according to respondents only 37 percent of the initiatives we learned about were targeted exclusively because of HEDIS measures, and only 6 percent were targeted exclusively because of CAHPS measures. Prioritizing quality-improvement activities is another way performance measures are used for targeting. Priorities may be established by comparing one's performance relative to external benchmarks or to internally defined expectations.

Evaluating Current Performance. The most frequently mentioned advantage of standardized performance measures is their utility for comparing performance against that of other plans. For example, one Medicaid plan in Washington made improving customer service and satisfaction its highest priority after its second consecutive year of poor performance on the CAHPS survey relative to other Medicaid plans in the state. Similarly, internally developed, "home-grown" measures, which may not be useful for external comparisons, can be used to assess the performance of different units within the same organization if the data are collected and reported consistently (e.g., plan-developed primary care physician satisfaction surveys).

Goal Setting. Plans also use performance measures to set improvement goals. Interestingly, when respondents were asked about the goals for each quality-improvement initiative discussed, they could provide a specific answer only 25 percent of the time. When specific goals were given they were usually determined by the results of benchmark performers in standardized data sets such as HEDIS and CAHPS. For example, one plan cited a 5 percent telephone call abandonment rate as the industry best-practice standard. For other measures plans strive for a simple percentage improvement on the plan's baseline rate (e.g., a 20 percent increase in the plan's baseline rate for usage of beta blockers after myocardial infarction).

Identifying the Root Cause. The measures that seem to be the most useful to plans for quality improvement are not necessarily those that are publicly

reported. Instead, plans frequently collect and analyze additional data to "drill down" to identify the root cause of substandard performance. For example, a plan may identify a low prenatal care utilization rate in their Medicaid population from reported measures but not necessarily know why the utilization rate is too low. Consistent with the conceptual model in Figure 1, plans often initiate studies to collect more detailed data from a variety of sources, such as provider surveys and member focus groups, to help identify the reason for the low utilization rate before designing quality-improvement initiatives.

Monitoring. Performance measures are also used to monitor the progress and outcomes of quality-improvement initiatives that have already been implemented. Sometimes the measures used for these purposes are the same as those that originally identified the need for the initiative. However, measures may also be developed to track an initiative's progress. The use of performance measures in this manner is consistent with our conceptual framework and allows plans to quantify achievement or to redesign initiatives that may not be having the intended result.

Range of Performance Measure Use: An Example

Although we identified five roles of performance measures, plans vary in their use of performance measures. We assessed range and variability by examining the different ways that plans use measures for similar initiatives. The most frequently cited initiatives pertained to diabetes care, with 11 plans describing quality-improvement activities related to this diagnosis. The scope of activity involved in the diabetes programs varied significantly from less aggressive initiatives that simply targeted members via educational materials to more aggressive initiatives that created disease registries and stratified diabetics by risk. Some plans took a comprehensive approach by focusing on both glycemic control by the patient and compliance with recommended guidelines by providers. Other initiatives focused on one aspect of appropriate diabetic care such as eye or foot exams.

Sixty-four percent of plans with diabetes initiatives cited HEDIS as a factor that led the plan to target diabetes, whereas 43 percent cited cost and high utilization of services (e.g., hospitalization because of infections, heart disease, etc.) associated with complications as the driving factor. Thirty-six percent indicated that the prevalence of diabetes was an important factor, whereas fewer respondents gave other reasons such as changes in NCQA and Health Care Financing Administration (HCFA) requirements (7 percent), variations in physician treatment patterns (7 percent), and findings from scientific studies or the published medical literature (14 percent). These

reasons are not mutually exclusive as some respondents listed multiple factors for their plan's decision to target diabetes care.

The most frequently mentioned source of data for evaluating current performance and setting goals was NCQA's *Quality Compass*. Until the recent addition of comprehensive diabetes measures, the HEDIS data set contained only one measure, the rate of retinopathy screenings.[5] Other information used to evaluate performance and set goals included data from published research studies and guidelines established by the American Diabetes Association. The association's guidelines recommend quarterly hemoglobin A1C screenings and annual eye and foot exams (Worrall, Freake, Kelland, et al. 1997; Stolar 1995).

However, only four respondents describing diabetes programs could list specific goals, which included reduction in hemoglobin A1C levels, increased eye and foot exams, and increased enrollment in the diabetes care management program. A range of data sources are used to identify the root cause of problems and to monitor progress. Five plans (36 percent) mentioned the HEDIS measures, four plans (29 percent) mentioned measures developed from administrative data (e.g., compliance with recommended appointments), three plans (21 percent) mentioned other internal measurement tools (e.g., data collected by nurse case managers), and two plans (14 percent) mentioned medical chart audits. One plan indicated that it used a special member-satisfaction-with-care survey to monitor the progress of its diabetes initiative.

Respondents cited the standardization of the HEDIS diabetes measures (29 percent) as one strength of the data set while citing a variety of weaknesses with these and other measures used to monitor diabetes care initiatives. Most of these weaknesses related to the quality of available data. For example, respondents mentioned problems with the accuracy of claims data, difficulty obtaining lab and pharmacy data, and the expense of obtaining data from medical chart reviews.

We performed similar analyses for two other programs: care for asthmatics and care for members with cardiovascular disease. Quality initiatives were implemented by several plans in both of these areas (eight plans for asthma, five plans for cardiovascular disease). These results are not reported because they yield conclusions similar to those for diabetes initiatives.

Characteristics of Useful Performance Measures

In addition to understanding the role that performance measures play in quality-improvement programs we sought to identify characteristics of measures that made them particularly useful. Our analysis suggests that the

most useful measures are standardized, actionable, timely, stable, capable of trending, measured at the appropriate unit of analysis, affordable and cost effective, and relevant.

Standardized. Standardization implies the establishment and acceptance of common criteria for the reporting of measures across units (e.g., plans, providers, clinics). The HEDIS criteria, for example, explicitly specify the eligible population, sampling methods, and algorithm used to calculate each measure. Standardization is most important for evaluating current performance and setting goals because it allows meaningful comparisons across health plans regionally and nationally or within units of the same organization. Insofar as targeting is based on the plan's performance relative to benchmarks, standardization is also important for this purpose. Standardization can be counterproductive, however, if it results in the reporting of measures that are not specific enough for plans to act upon.

Actionable. Numerous respondents used the term "actionable" to describe useful measures and to explain why some measures were less useful for quality improvement. By actionable, respondents meant that the measure helps pinpoint the changes a plan must undertake to realize an improvement. For example, a high telephone abandonment rate may suggest that the plan should hire more customer service representatives, whereas the action required to improve other measures, such as a low rate of follow-up after inpatient mental health hospitalization, may not be immediately clear. This characteristic is most important for identifying the root cause of problems after they have been targeted.

Timely. Several respondents indicated the need to make decisions based on current data. However, for current public reporting systems such as HEDIS there can be as much as an 18-month lag between data entry on the medical chart and benchmark data reporting by NCQA in its *Quality Compass* product. Some plans find this lag difficult for targeting, evaluating current performance, and setting goals, although it should be noted that little is lost by identifying benchmark performers from older data. The more pressing issue seems to be a significant lag between the actual clinical encounter and the time that HEDIS data are collected and reported by many plans, making it difficult to have discussions with particular providers about specific cases. However, not all plans have this problem as some are more adept than others at collecting and analyzing data closer to the actual encounter. One area in which plans do appear to be collecting data in real time is customer service. Many plans have automated phone-tracking systems that produce routine reports such as speed of answer, complaints, and call transfers.

Stable and Capable of Trending. Numerous respondents expressed frustration that publicly reported measures change frequently, increasing the organizational burden of reporting and reducing the longitudinal comparability of results. Because collection of most clinical measures requires computer programming or medical chart review, both of which can be costly, the need to re-educate a plan's workforce about changes in measure-reporting methodology is burdensome. Similarly, some respondents argued that it is difficult to accurately trend measures when they are constantly changing. Several respondents who were particularly knowledgeable about the techniques of continuous quality improvement viewed stability in a slightly different manner. These respondents agreed that having a stable set of measures is beneficial as long as they focus on problems or issues the organization needs to address (e.g., how to increase compliance with recommended mammography screenings). However, once the problem has been resolved, these respondents believe that resources can be used more efficiently by moving on to other measures.

Appropriate Unit of Analysis. Several respondents described a lack of agreement between a standardized measure's unit of analysis, typically the health plan, and the locus of care or service provision, typically the individual provider, clinic, hospital, or plan customer service department. This issue is very much related to actionability because many plans find it difficult to identify which provider or clinic is the source of problems. Having only aggregate, plan-level data is limiting for purposes of targeting, identifying the root cause of problems, and monitoring. Even if the problem provider or clinic is known to the plan, providers may deny problems until they see data specific to their practice rather than data aggregated across multiple physicians and practices.

Affordable and Cost-Effective. Respondents identified useful measures as those leading to improvements in quality at low cost. However, few respondents could estimate the aggregate cost of performance measurement in per-member per-month terms or as a percentage of administrative costs, not to mention the cost of specific measures. The irony is that although respondents say that they want cost-effective measures, they cannot determine whether a measure is cost-effective. Instead, it appears that respondents were expressing an interest in quality initiatives that improve the bottom line in terms of both patient benefit and cost reduction regardless of the cost of the measures involved.

The concern about cost may be exacerbated by the belief that performance measurement itself is not valued in the marketplace by purchasers or

consumers. Many respondents believe that purchasers and consumers base their contracting and enrollment decisions on price alone; therefore, plans do not benefit by measuring for measurement's sake or by outperforming their competitors on quality. When pressed on this issue many respondents agreed that some purchasers use measurement and accreditation results to establish minimum acceptable thresholds of quality and then purchase based on price. However, this view was not unanimous.

Relevant. For a measure to be useful it must be relevant to the population a plan serves. Some respondents suggested that measures are irrelevant if the plan is already performing satisfactorily on the measure in question. Measures were also described as irrelevant if they were not germane to the population served. For example, one medical director mentioned the decision to incorporate into HEDIS a measure of the percentage of women in a plan receiving chlamydia screenings. Although the respondent acknowledged that this measure had clinical merit for certain women, most notably those enrolled in Medicaid, he felt that the measure was inappropriate for the commercial population that his plan served. Interestingly, a reviewer used this comment to point out that one of the advantages of externally required measurement may be its ability to define relevancy. For example, the reviewer noted that, "The comment on chlamydia reflects a common misperception among the medical community and one that underscores the benefits of some of the standardized measurement activities. The measure is consistent with the U.S. Public Health Service guidelines on preventive care, specifically that women in the age range of the measure are at high risk (not just women on Medicaid)."

Strengths and Weaknesses of HEDIS and CAHPS for Quality Improvement

Because HEDIS and CAHPS are so prominent in today's managed care landscape, we asked plan respondents to use a five-point scale (five being the best) to separately answer two questions about HEDIS and CAHPS. First, we asked respondents, "How well do the data represent what they attempt to measure?" Second, we asked, "How useful is HEDIS/CAHPS to your plan for purposes of quality improvement?" The results from these questions are reported in Table 4, but readers are cautioned that the cell sizes are very small. Respondents were generally in the middle of the scale for both questions, with a mean score of 3.35 and 3.60, respectively, for HEDIS and a mean score of 3.21 and 3.13, respectively, for CAHPS. When the results are examined by position they are consistent for medical directors and quality improvement directors but generally lower for CEOs. The lower scores provided by CEOs

may reflect the fact that CEOs are not as actively involved in the collection and use of HEDIS and CAHPS measures compared to the other respondents. Although the results vary by state, there is not a discernible pattern across the four questions. The greatest amount of variation pertains to the utility of CAHPS for quality improvement, which had a mean rating of 2.88 in Kansas and a mean rating of 3.63 in Maryland.

Subsequent discussion with respondents indicated that the most frequent use of both HEDIS and CAHPS is for evaluating current performance and setting goals. The utility of these measurement systems for these purposes results primarily from the fact that they are both used nationally and comparative results can be obtained from NCQA's *Quality Compass*. Respondents generally cited HEDIS as being more useful than CAHPS for targeting and monitoring. Most plans, even the three that have no intention of becoming accredited by the NCQA, collect HEDIS data and have specific quality-improvement initiatives to address one or more of the measures contained in HEDIS.

Several disadvantages of HEDIS were noted. First, 17 percent of the respondents indicated that collecting and reporting HEDIS measures is costly, particularly verification of provider credentials and chart reviews for annual measure reporting.[6] Second, several respondents questioned the comparabil-

Table 4: Ratings of HEDIS and CAHPS

	HEDIS Mean Accuracy Rating (1–5)	*HEDIS Mean Utility Rating (1–5)*	*CAHPS Mean Accuracy Rating (1–5)*	*CAHPS Mean Utility Rating (1–5)*
Overall Mean Ratings	3.35 ($n = 34$)	3.60 ($n = 34$)	3.21 ($n = 33$)	3.13 ($n = 32$)
Mean Ratings by Position				
Medical director	3.35 ($n = 12$)	3.56 ($n = 12$)	3.21 ($n = 10$)	3.13 ($n = 10$)
CEO	3.16 ($n = 5$)	3.34 ($n = 5$)	2.90 ($n = 5$)	2.60 ($n = 5$)
Quality improvement manager	3.40 ($n = 17$)	3.53 ($n = 17$)	3.22 ($n = 18$)	3.32 ($n = 17$)
Mean Ratings by State				
Kansas	3.20 ($n = 9$)	3.28 ($n = 9$)	3.00 ($n = 8$)	2.88 ($n = 8$)
Maryland	3.19 ($n = 8$)	3.63 ($n = 8$)	3.33 ($n = 9$)	3.63 ($n = 8$)
Pennsylvania	3.75 ($n = 9$)	4.02 ($n = 9$)	3.17 ($n = 9$)	3.03 ($n = 9$)
Washington	3.25 ($n = 8$)	3.31 ($n = 8$)	3.36 ($n = 7$)	3.00 ($n = 7$)

Note: The HEDIS and CAHPS accuracy ratings were based on the question, "How well do the data represent what they attempt to measure?" Similarly, the HEDIS and CAHPS utility ratings were based on the question, "How useful is HEDIS/CAHPS to your plan for purposes of quality improvement?" Respondents were asked to answer on a scale of 1 to 5 (5 being the best).

ity of HEDIS data between different plan models, such as independent practice association plans and group model plans, because of variations in plan-provider relationships. Third, several respondents suggested that HEDIS's continuous enrollment criteria are problematic for Medicaid plans because there is significant turnover in this population.

After years of using a variety of membership and satisfaction surveys most plans are now adopting CAHPS. Similar to HEDIS, the most frequently reported advantage of CAHPS is for targeting, evaluating current performance, and goal setting. However, as reported by the study respondents, the utility of CAHPS for quality improvement is less established for several reasons. First, 50 percent of the respondents indicated that the survey questions are not specific enough to identify the locus of problems. For example, one respondent cited the question, "Were you treated with respect?" as an example of a measure that is not actionable because the source of the problem (i.e., which provider or clinic) is unknown. Even more specific questions, such as those pertaining to ease of appointment access, are hard to act upon because the measure is based on a random sample of plan members, each of whom may schedule appointments with different providers. Hence, action requires detailed root-cause analysis, but such analysis is difficult because survey respondents are anonymous to the plan. Second, 31 percent of the respondent plans already conduct their own primary care physician–based member surveys, in some cases as frequently as quarterly. Respondents from plans using these surveys claim that they are more beneficial for improving quality and solving care and service problems because the unit of analysis reflects the locus of care. Third, several respondents suggested that the cost of CAHPS is too high and response rates are too low. Two plans indicated that they would not conduct CAHPS if they were not required by NCQA and HCFA.

DISCUSSION

Although the results of our study cannot be generalized beyond our sample of MCOs, they can be used to structure research questions and to develop hypotheses for future research that is generalizable to a broader population of managed care plans. For the MCOs in our study we found the following answers to our research questions.

- They are engaged in measurement and, consistent with our conceptual model, there is a link between measurement and improvement.

However, there also seems to be significant variation in how these plans use performance measures, even for similar initiatives, but the reasons for this apparent variation could not be identified in our study.

- Performance measures seem to play at least five roles for quality-improvement activities, although specific measures seldom serve all roles. Measures can be used to evaluate current performance, target quality-improvement initiatives, establish new performance goals, identify the root cause of problems, and monitor the progress of quality-improvement initiatives.
- Respondents identified several characteristics of performance measures (e.g., standardized, actionable) that make them particularly useful for quality-improvement activities and that developers of performance measures might want to consider.
- HEDIS and CAHPS are the most widely used sets of performance measures and are particularly useful to plans for targeting, evaluating performance relative to other plans, and establishing quality-improvement goals. However, respondents also identified specific issues that compromise the utility of these standardized data sets for quality-improvement programs.

In addition to assessing the generalizability of these principal findings future research might address several additional questions. For example, because there does appear to be variation among plans in terms of measurement and improvement capabilities, even in our small sample, future studies might identify the reason for this variation. For example, does the ability to successfully engage in these activities depend on a plan's profit status, leadership, experience, market environment, or type of relationship between the insurance plan and the providers of care (i.e., plan model type)? Systematically identifying characteristics of MCOs with more- and less-effective performance measurement and quality improvement may go a long way toward explaining these variations.

Another important question centers on the value of standardized performance measures being required by external stakeholders. There is considerable debate about whether requiring plans to report standardized performance measures enhances or hinders quality-improvement efforts. Advocates argue that requiring standardized collection and reporting provides quality with a seat at the table it might not otherwise have. Opponents contend that standardized measurement unnecessarily limits the focus of quality-improvement activities, replacing the flexibility required for true continuous quality improvement with a one-size-fits-all approach. Future research might

advance this debate by systematically assessing the effect of standardized measurement on the effectiveness and breadth of quality-improvement efforts and if necessary devising ways to reduce potentially negative consequences.

Another question worth asking is whether the marketplace is providing sufficient incentives to prompt plans to take performance measurement seriously. Stated differently, if measurement is in fact related to improvement, as many believe, and if purchasers value quality, then why are quality measurement and quality improvement still in their infancy in MCOs? The answer to this question likely depends on answers to related questions, for example, how purchasers weigh cost versus quality, the comparative costs and benefits of measuring performance, and variations in the perceived benefit by stakeholders.

Finally, there is the question of the most cost-effective configuration of performance measurement. A number of stakeholders, including the plans themselves, purchasers, consumers, and regulators, are currently pressing for performance measurement, but each for their own purposes. Performance measurement, however, can be a costly exercise. Thus, it becomes important to learn how the performance-measurement dollar can be allocated both cost effectively and in a way that meets a diversity of needs. Our research is clearly unable to answer these questions but hopefully will set the stage for future studies that can.

ACKNOWLEDGMENTS

The authors are grateful for the participation of the study respondents from 24 health plans in four states and pilot study respondents from four plans in New Jersey. We also appreciate the valuable comments provided by Ron Hays and Pamela Farley Short.

NOTES

1. Kansas and Washington were demonstration sites for evaluating the CAHPS survey and its reports. The state of Maryland issued a performance report for commercial plans that was based on CAHPS and data from HEDIS. In Pennsylvania the Hospital and Health System Association produced a comparative HEDIS report. The Medicare risk plans in all states are the subject of a CAHPS report sponsored by HCFA.
2. Fewer than three initiatives were discussed with some respondents because of time constraints.

3. After reviewing the interview notes the study authors independently selected plans that were significant users of performance measures for quality-improvement purposes. This list was then narrowed by majority agreement.
4. To calculate this statistic we defined a performance measure as any data collected and used by the plan regardless of whether the data were externally reported. For example, a plan may have analyzed patterns in its claims data for reasons not related to external reporting. If such an analysis identified an area to be targeted, perhaps because of high expenditures, we attributed the initiative to a performance measure.
5. Now HEDIS contains not only retinopathy screenings but also glycohemoglobin testing and control, lipid level testing and control, and regular screening for kidney disease (NCQA 1999b).
6. NCQA recently announced a system of rotation so that MCOs may use certain measurement results for two years (NCQA 1999c), possibly minimizing the burden of reporting.

REFERENCES

Carman, K. L., P. F. Short, D. O. Farley, J. A. Schnaier, D. B. Elliott, and P. M. Gallagher. 1999. "Early Lessons from CAHPS Demonstrations and Evaluations." *Medical Care* 37 (3): 97–105.

Chernew, M., and D. P. Scanlon. 1998. "Health Plan Report Cards and Insurance Choice." *Inquiry* 35 (Spring): 9–22.

Coltin, K. L., and D. B. Aronow. 1995. "Quality Assurance and Quality Improvement in the Information Age." In *Quality in Health Care,* edited by N. O. Graham, 223–48. Gaithersburg, MD: Aspen Publishers.

Gabel, J. R., K. A. Hunt, and K. Hurst. 1998. *When Employers Choose Health Plans: Do NCQA Accreditation and HEDIS Data Count?* New York: KPMG Peat Marwick, LLP.

Goverman, I. L. 1994. "Orienting Health Care Information Systems Toward Quality: How Group Health Cooperative of Puget Sound Did It." *The Joint Commission Journal on Quality Improvement* 20 (11): 595–605.

Health Care Financing Administration (HCFA). 1997. "Medicaid National Summary Statistics: National Summary of Medicaid Managed Care Programs and Enrollment." [Online.] http://www.hcfa.gov/medicaid/trends1.htm.

———. 1999. "Penetration Rates from 1993–1998: National Summary Table." [Online.] http://www.hcfa.gov/medicaid/trends98.htm.

———. 2000a. "Medicare Managed Care Market Penetration for All Medicare Plan Contractors-Quarterly State/County Data Files (March 2000 Data)." [Online.] http://www.hcfa.gov/medicare/mpsct1.htm.

———. 2000b. "Medicaid Managed Care State Enrollment–June 1999." [Online.] http://www.hcfa.gov/medicare/mcsten99.htm.

Hillman, A. L., and N. Goldfarb. 1995. "Exemplary Quality Improvement Programs in HMOs." *The Joint Commission Journal on Quality Improvement* 21 (9): 457–64.

The InterStudy Competitive Edge: HMO Industry Report. 1999. Biannual Report 9 (1). St. Paul, MN: InterStudy.

Jaeger, B. J., A. D. Kaluzny, and C. P. McLaughlin. 1994. "TQM/CQI: From Industry to Health Care." In *Continuous Quality Improvement in Health Care,* edited by C. P. McLaughlin and A. D. Kaluzny, 11–32. Gaithersburg, MD: Aspen Publishers.

Kaiser Family Foundation. 1999. *Medicare Managed Care Fact Sheet.* [Online article; retrieved September 1999] http://www.kff.org/content/archive/2052/MedicareManaged.pdf. Washington, DC: Kaiser Family Foundation.

Kinney, C. F., and R. G. Gift. 1997. "Building a Framework for Multiple Improvement Initiatives." *The Joint Commission Journal on Quality Improvement* 23 (8): 407–23.

KPMG Peat Marwick. 1998. *Health Benefits in 1998.* Arlington, VA: KPMG.

Landon, B. E., and A. M. Epstein. 1999. "Quality Management Practices in Medicaid Managed Care." *Journal of the American Medical Association* 282 (18): 1769–75.

Landon, B. E., C. Tobias, and A. M. Epstein. 1998. "Quality Management by State Medicaid Agencies Converting to Managed Care." *Journal of the American Medical Association* 279 (3): 211–16.

McLaughlin, C. P, and A. D. Kaluzny. 1994. "Defining Total Quality Management/ Continuous Quality Improvement." In *Continuous Quality Improvement in Health Care,* edited by C. P. McLaughlin and A. D. Kaluzny, 3–10. Gaithersburg, MD: Aspen Publishers.

National Committee for Quality Assurance (NCQA). 1995. *HEDIS 2.5: Updated Specifications for HEDIS 2.0.* Washington, DC: NCQA.

———. 1999a. *Quality Profiles.* Washington, DC: NCQA.

———. 1999b. "HEDIS 1999-Update to Volume 1: Narrative (March)." [Online.] http://www.ncqa.org/indexsm.htm. Washington, DC: NCQA.

———. (P. Torda.) 1999c. "Memorandum: Rotation Strategy for HEDIS 2000." [Online.] ftp://www.ncqa.org/docs/hedis/rotate1.doc. Washington, DC: NCQA.

———. 2000a. "NCQA Programs." [Online.] http://www.ncqa.org/Pages/Main/programs.htm. Washington, DC: NCQA.

———. 2000b. "NCQA's Health Plan Report Card." [Online. Retrieval 8/25/00.] http://www.ncqa.org/pages/hprc/index.html. Washington, DC: NCQA.

Peterson, M. A. 1999. "The Managed Care Backlash-Introduction: Politics, Misperceptions, or Apropos?" *Journal of Health Politics, Policy, and Law* 24 (5): 873–86.

Scanlon, D. P., and M. E. Chernew. 1999. "HEDIS Measures and Managed Care Enrollment." *Medical Care Research and Review* 56 (Suppl. 2): 60–84.

Scanlon, D. P., and T. J. Hendrix. 1998. "Health Plan Accreditation: NCQA, JCAHO, or Both?" *Managed Care Quarterly* 6 (4): 52–61.

Scanlon, D. P., E. Rolph, C. Darby, and H. E. Doty. 2000. "Are Managed Care Plans Organizing for Quality?" *Medical Care Research and Review* 57 (Suppl. 2): 9–32.

Stolar, M. W. 1995. "Clinical Management of the NIDDM Patient." *Diabetes Care* 18 (5): 701–07.

Worrall, G., D. Freake, J. Kelland, A. Pickle, and T. Keenan. 1997. "Care of Patients with Type II Diabetes: A Study of Family Physicians' Compliance with Clinical Practice Guidelines." *The Journal of Family Practice* 44 (4): 374–81.