

MANAGED CARE AND PERFORMANCE MEASUREMENT: IMPLICATIONS FOR INSURANCE MARKETS

Dennis P. Scanlon* and Michael Chernew†

ABSTRACT

Performance measurement systems and *report cards*, which attempt to measure and report the quality of care provided by managed health-care organizations, have become mainstream in health insurance markets as managed care penetration continues to increase. However, little is known about the impact formal plan evaluations have on the contracting and enrollment decisions made by health insurance purchasers and consumers. Information regarding the link between performance evaluations and enrollment is crucial for those charged with projecting future enrollments in and risk profiles of managed care organizations. This paper describes the performance measurement systems currently being used to evaluate managed care plans and reviews the empirical literature for evidence regarding the impact of measures on plan enrollments.

1. HEALTH PLAN PERFORMANCE MEASUREMENT: WHAT IS IT AND WHY IS IT IMPORTANT?

The health insurance market has undergone major changes in the past decade as enrollment in traditional indemnity plans has decreased in favor of health maintenance organizations, preferred provider organizations, and other forms of managed care (MC) plans. Broadly defined, MC refers to organizational entities that contractually agree to both finance and arrange for the delivery of health care for a defined population of enrollees.

Because traditional indemnity plans placed few restrictions on who could provide health care, individual insureds ultimately retained some responsibility, through selection of provider at the time of illness, for ensuring that quality care was received. Similarly, because many MC plans limit the availability of providers (physicians and hospitals) by virtue of their contract-

ing decisions, MC plans can affect the quality of care provided.

The performance of an MC plan, including the quality of care delivered, will reflect the composition of its provider network, the financial incentives placed on providers, and the administrative controls and educational programs designed to influence care. Plans vary widely in the breadth of their provider networks, the processes used to select physicians and hospitals into their networks, the financial and nonfinancial barriers associated with receiving care outside of the network of providers, and the extent to which financial risk is passed on to providers.

Utilization review, clinical practice guidelines, and educational initiatives for management of specific diseases can also influence the practice patterns of individual physicians. Many of these programs rely on sophisticated use of information technology. All of these activities can influence the performance of MC plans and the quality of care received by their enrollees.

Because MC providers have a strong incentive to control costs, and because purchasers increasingly want to reduce payments to MC plans, a natural concern is the impact that such incentives and cost reductions will have on quality of care. Reduced costs often entail reductions in utilization, which can affect quality. Performance measurement, to the extent that

*Dennis P. Scanlon, Ph.D., is Assistant Professor in the Department of Health Policy & Administration, Center for Health Policy Research, The Pennsylvania State University, 116 Henderson, University Park, PA 16802-6500, e-mail, dpscanlon@psu.edu.

†Michael Chernew, Ph.D., is Associate Professor in the Departments of Health Management & Policy, Economics, and Internal Medicine, University of Michigan, SPH II, 1420 Observatory Street, Ann Arbor, MI 48109-2029, e-mail, mchernew@umich.edu.

it measures quality of care, can be used to monitor quality in enrolled populations. Unfortunately, current performance measurement systems do not measure many important aspects of quality of care (Mainous and Talbert 1998). Nevertheless, the growth in MC enrollment, from 29% in 1988 to 86% in 1998 (KPMG 1998), has been paralleled by an increased emphasis on the measurement and evaluation of the performance and quality of care provided by MC organizations. Performance measurement systems are increasingly striving to better measure quality.

Many analysts have suggested that public dissemination of comparative plan performance information, typically in the form of ratings and rankings, would improve market outcomes by encouraging consumers to become more competitive “shoppers” and to consider plan attributes other than price (Enthoven 1988, 1989). Proponents of this approach believe that well-informed consumers will “vote with their feet” and ultimately drive inferior quality plans from the marketplace through disenrollment while rewarding superior quality plans through increased enrollment. In theory, such a marketplace of savvy consumers could have a profound impact on a plan’s enrollment mix, average level of risk, medical loss ratio, and other important actuarial characteristics (Pyenson et al. 1998). Unfortunately, there is currently little evidence that consumers use such information in the manner hypothesized.

This paper discusses performance measurement systems and current empirical evidence regarding the impact of performance measurement on health plan enrollment. Section 2 discusses the history of performance measures in medical care and describes measurement systems that are currently being utilized in MC. Section 3 discusses crucial methodological issues for examining the impact of performance measures on enrollment. Section 4 reviews the current empirical evidence regarding the relationship between performance measures and enrollment. Section 5 concludes with a discussion of important policy issues.

2. HISTORY OF PERFORMANCE MEASUREMENT

Although performance measurement in medical care has received a great deal of media attention recently, efforts to evaluate the quality of health care have existed for several centuries. The earliest documented effort to evaluate quality of care is attributed to Florence Nightingale in the nineteenth century. Nightin-

gale, a nurse, identified differences in mortality statistics among British hospitals and argued that research was needed to explain why results differed and how hospitals could change their processes to reduce mortality (McDonald 1998, Iezzoni 1996).

Earnest Codman, a surgeon at the Massachusetts General Hospital in the early twentieth century, is considered the father of performance measurement in the U.S. Like Nightingale, Codman’s work focused on the hospital. He is best known for creating a formal framework for evaluating hospital quality, which he labeled “the end result idea.” Codman suggested that hospitals should monitor patients during their inpatient stay and for a period of time following their discharge. As part of the evaluation process, Codman proposed collecting data on key outcome indicators such as mortality and morbidity, so that hospital performance, or the end result, could be monitored. Codman believed that in the case of outcomes such as death or excess loss of function, hospitals should retrospectively identify the cause by systematically examining the entire process of care. For example, the unexpected death of a surgical patient should be examined and explained in terms of mistakes by the surgeon, poor nursing care, improper patient compliance with discharge orders, or simply chance. Codman believed that hospitals and providers could only improve the quality of care provided by identifying the reasons for bad outcomes (Donabedian 1989, McLendon 1990).

Avedis Donabedian, a physician and self-proclaimed disciple of Codman, expanded Codman’s work and is widely known for developing a general model identifying categories of measures that can be used to monitor the quality of care given by any health-care provider or in any health-care environment. Specifically, Donabedian suggested that in the absence of true measures of quality, measures of structure, process, and outcome can be used in isolation or in combination to evaluate the quality of care provided (Donabedian 1988). Donabedian defined *structural measures* as those characteristics of the provider or institution thought to be correlated with high-quality care; for example, proper education, licensure, training for physicians and nurses, and appropriate staffing and technological capacity for hospitals. *Process measures* reflect the manner in which care is provided and include tests, treatments, or procedures thought to be beneficial for a certain illness or disease. For example, evidence indicates that prescription of beta-blockers following myocardial infarction is efficacious. A specific process measure can capture the percentage of

a physician's or hospital's heart attack patients that receive a prescription for beta-blockers following a heart attack. Finally, outcome measures are similar to the "end result" proposed by Codman. Namely, what happens to the patient at the end of the day? Does the patient live? Has the patient's physical functioning been restored?

Donabedian suggested that the choice of measures should depend on the goal of the quality assessment study and data limitations (Donabedian 1988). In addition, specific types of measures can be used in isolation or several measures can be used in tandem to judge the quality of care provided. For example, the use of outcome measures, such as surgical death rates, in isolation does not indicate that poor or good quality care was provided, unless the measure is adjusted for the illness severity of the patient population (that is, case-mix) and the sample size is sufficient. In other words, while it is possible that poor outcomes are due to inferior care, they can also stem from a systemically sicker patient base or from chance. Similarly, structural and process measures do not guarantee attainment of a positive outcome. Rather, the presence of appropriate structure and process should be correlated with positive outcomes.

Current performance measurement efforts build on the decades of work by the scholars identified above as well as others who have researched appropriate methods to measure quality and compare performance across organizations and providers. One such effort is the Health Plan Employer Data and Information Set (HEDIS), which is the most widely used system for measuring and comparing the performance of care provided by MC plans. HEDIS was created in 1989 as an evaluation tool to be used for quality improvement by a consortium of large employers (purchasers) and health plans. Since its birth, HEDIS has undergone several revisions (currently HEDIS 2000), and consists of about 100 measures (National Committee for Quality Assurance 1997). The measures contained in HEDIS are largely process measures (as defined above), but structural and outcome measures are also included. For example, mammography rates for women aged 52–69 is a process measure included in the data set, while percentage of low-birthweight babies is an outcome measure, and percentage of physicians with board certification is a structural measure.

The popularity of HEDIS has grown due to competitive market forces and industry regulation. Compliance has been significant and primarily driven by large purchasers (such as corporations, state Medicaid pro-

grams, and so on) who demand that plans report HEDIS data as a requirement for contracting. In addition, many purchasers have also required plans to be accredited by the National Committee for Quality Assurance (NCQA), the primary accrediting body for the MC industry. Prior to 1999, the NCQA required, as part of its accreditation process, that plans demonstrate the existence of a quality measurement and improvement program. Since HEDIS is administered by NCQA, most plans used HEDIS to demonstrate quality measurement and to document quality improvement, even though use of HEDIS was not itself a requirement for accreditation. However, since 1999 sixteen HEDIS measures have been required as part of the NCQA accreditation process, leaving plans with little choice but to report at least a portion of the HEDIS dataset if they desire NCQA accreditation (NCQA 1999).

A more recent health plan performance measurement tool is the Consumer Assessment of Health Plan Survey (CAHPS), sponsored by the Agency for Health Care Policy and Research. Unlike HEDIS, which relies primarily on administrative records documenting clinical procedures, CAHPS uses survey methodology to capture the opinions of health plan enrollees about experiences with their plans. CAHPS measures inform consumers about how others rate the care provided by their health plans. Like HEDIS, CAHPS measures have been disseminated publicly and have been incorporated into plan report cards. Since 1999, the CAHPS survey has been incorporated into the HEDIS dataset, resulting in an assessment tool that includes administrative, clinical, and member survey data.

The growth and popularity of performance measures has resulted in widespread public dissemination of results and the release of health plan ratings and rankings based on these measures. Public dissemination of HEDIS measures has been controversial because the dataset was originally designed for internal use by plans and purchasers, not public use, and because HEDIS contains methodological limitations (such as lack of case-mix adjustment) that make valid comparisons between plans difficult. Nonetheless, many large purchasers, government agencies, and periodicals have disseminated plan ratings and rankings to consumers based on the HEDIS or CAHPS datasets. Most of these initiatives are pursued to inform consumers and to facilitate competition between plans on non-price attributes.

The public dissemination of plan ratings has potentially profound effects for health insurance markets. If consumers value and use the information provided by

these ratings, health plans may witness shifts in enrollment based on their performance relative to their competition. Widespread shifts in enrollment can affect the financial solvency and operational capability of health plans. Hence, information about the impact of ratings on plan enrollment is crucial, especially for those charged with projecting plan financial characteristics (Pyenson et al. 1998).

3. IMPORTANT METHODOLOGICAL ISSUES FOR EXAMINING THE RELATIONSHIP BETWEEN PERFORMANCE MEASURES AND PLAN ENROLLMENT

Three methodological issues make the impact of performance measures on plan enrollment in health insurance markets difficult to study. Before discussion of these issues, several terms must be defined. *Report card* is a term used to describe any initiative to provide comparative plan-performance information, regardless of whether the provided information includes individual or aggregated measures. *Measures* refer to specific raw measures such as those contained in HEDIS or CAHPS. *Indices* refer to aggregations of two or more individual measures. For example, many report cards combine the HEDIS preventive care measures (such as mammography screening rates, childhood immunization rates, and so on) to obtain a prevention index. *Ratings* involve normative judgments about a plan's performance, or its performance relative to other plans, based on individual measures or indices. For example, ratings such as "above average," "needs improvement," or "superior" are commonly assigned to health plans based on the results of individual HEDIS measures or HEDIS-based indices.

The first methodological issue is whether to examine the impact of individual performance measures or aggregated measures in the form of ratings, rankings, or indices. Choosing whether to study health plan enrollment in relation to measures, indices, or ratings is difficult. The studies available in the literature to date have generally differed in their approach to this issue. Focus group studies have asked individuals to list specific measures that would be important in selecting a health plan. Other studies have provided individuals with specific measures for hypothetical choice offerings and have asked individuals to choose a health plan based on the provided information. Still other studies have examined the actual plan choices of employees who received plan ratings for each option available to them. This last class of studies statistically

models plan choice as a function of the provided plan ratings, which allows researchers to identify correlations between specific ratings and enrollment but does not indicate which measures comprising the ratings are the most influential in plan selection.

The second important methodological issue is the source of data used in studies examining the impact of performance measures and ratings on enrollment decisions. Typically, the data used is obtained from consumer surveys, focus groups, laboratory experiments, or "revealed preference" studies that rely on administrative data to document actual enrollment decisions. Consumer surveys offer the advantage of asking specific questions about the decision to enroll in one's current health plan. Although this methodology is direct, it may suffer from recall bias on the part of the respondent. Focus group studies allow the researcher to probe issues discussed by participants. Laboratory experiments require subjects to make hypothetical enrollment decisions from a set of health plan offerings with various levels of plan attributes (for example, price, measures, indices, ratings, and so on). Plan choice is then monitored to identify which plan characteristics had the most influence on plan choice. However, all these types of studies suffer from potential bias because consumers are not making binding choices, and therefore the results may not be generalizable. Revealed preference studies are advantageous because they utilize administrative data documenting the actual choices made by consumers. Combined with information about the available choice options and the attributes of the available choices, these studies avoid the binding-choice problems associated with surveys, focus groups, and laboratory experiments, although they generally include less qualitative information about the choice process.

The third methodological issue is exclusion of information thought to influence the enrollment decision in studies examining the importance of performance measures and ratings. Revealed preference studies suffer from the presence of important plan attributes that are unobserved by the researcher. For example, analysts believe that the most important determinant of health plan enrollment is plan participation by one's current physician. Unfortunately, no revealed preference studies observe sufficient information on plans, physicians, or consumers to control for this variable empirically. Other variables, such as a plan's ease of specialty referral or benefit structure, may also be important predictors of plan enrollment but are usually unobserved by the researcher. The absence of such information is less problematic in lon-

gitudinal revealed preference studies if the unobserved variables can be assumed to remain constant from year to year. However, when analyzing cross-sectional data, researchers typically assume that unobserved attributes are independently and identically distributed across observations. If this assumption is correct, unbiased estimates of the impact of the observed variables on plan choice can be estimated.

4. CURRENT KNOWLEDGE ABOUT THE RELATIONSHIP BETWEEN MEASURES AND RATINGS AND PLAN ENROLLMENT

Studies were chosen for this review by conducting thorough literature searches using MEDLINE and HEALTHSTAR. Studies were also identified for inclusion from the bibliographies of studies obtained from the literature searches. One unpublished study was included (Farley et al. 1999) because it was one of only three revealed preference studies available. Table 1 lists the studies that were reviewed for this paper, along with information on the focus, sample, and methodology used. They will be discussed according to the methodology used to examine the importance of comparative health plan information. Studies analyzing data obtained from survey, focus group, laboratory, or case study methodology will be discussed separately from revealed preference studies.

Findings from Surveys, Focus Groups, Laboratory Experiments, and Case Studies

The majority of current published and unpublished studies examining consumer use of health plan report cards analyze data obtained from surveys, focus groups, laboratory experiments, and case studies. Three issues make interpretation of findings from these studies difficult. First, the results are often reported qualitatively, making comparisons between studies difficult. Second, although many studies use similar methods—such as focus groups or surveys—the interview protocols differ between studies, making comparisons difficult. Third, because these studies do not use data obtained from binding enrollment decisions, it is not clear how reliable the results are. McGee and Knutson (1994) emphasize the importance of this third issue:

Fundamentally, it is questionable how much insight can be gained into the real-life impact of compara-

tive quality information by asking people what impact they *think* information would have on them, particularly if situational factors are highly influential. People may also tend to answer hypothetical situations more in terms of what they think they *ought* to do (pay attention to comparative information) than in terms of what they probably actually *would* do (which is hard or impossible to know). Consequently, studies based on hypothetical situations may tend to overestimate the extent to which consumers will use comparative information. (p. 11)

These caveats aside, current studies in this category provide some insight into consumer views regarding the usefulness of plan report cards and measures. Most of the studies find that consumers desire comparative information, although few studies find evidence that consumers with access to such information actually use it in the decision-making process. For example, Robinson and Brodie (1997) report that 87% of those who saw comparative information thought it would be useful in enrollment decisions, but only 34% of these individuals reported actually using the information when deciding which plan to enroll in.

Hibbard and Jewett (1996) found similar inconsistencies between what consumers say and what they do. The authors conducted focus groups to assess the usefulness and salience of the quality-of-care information currently being provided in health plan report cards. Respondents indicated a general interest in quality measures and a specific interest in preventive care and satisfaction measures relative to undesirable event rates, such as nosocomial infections or measures of disciplinary action against the plan. However, when asked to make a hypothetical choice between two plans, one performing better on prevention and satisfaction and the other on undesirable event rates, consumers overwhelmingly selected the plan with better results on undesirable event rates. The authors believe the inconsistency arises because consumers, when making decisions, give more preference to measures they cannot control (that is, undesirable events) than to measures they can control (that is, utilization of preventive care services).

When asked about the sources of health plan information that are most commonly used and trusted, respondents from several studies indicated that friends and family were very important sources. For example, 57% of respondents in the Robinson and Brodie study indicated that family and friends influenced their choice of health plan. In many of the studies, employers, plans, and the government are viewed skeptically as sources of information, while responses are mixed

**Table 1
Studies Reviewed**

Authors	Year Published	Methodology	Sample	Focus
Spranca et al.	2000	Laboratory experiment/ hypothetical plan choice	311 privately insured adults from Los Angeles County, California	Affect of CAHPS reports on health plan selection
Booske et al.	1999	Study of hypothetical plan choice	Wisconsin state employees	Eliciting consumer preferences for health plans
Farley et al.	1999	Analysis of actual health plan enrollment data	New Medicaid enrollees choosing MC coverage in New Jersey in 1998	Impact of reported CAHPS data on MC enrollment
Scanlon and Chernew	1999	Market-share regression analysis	National sample of employees choosing family MC coverage in 1996 through a Fortune 50 company	Impact of health plan ratings and HEDIS measures on MC plan enrollment
Wicks et al.	1999	Case studies and focus groups	5 public and private health plan purchasing initiatives	Consumer use and attitudes about comparative health plan information
National Committee for Quality Assurance	1998	Survey analysis	Sample of consumers in Denver and St. Louis	Consumer use of and reaction to comparative health plan performance information
Knutson et al.	1998	Survey analysis	Sample of two groups of Minnesota state employees	Consumer use of a comparative health plan report card
Gabel et al.	1998	Survey analysis	1,502 randomly selected public and private employers with at least 200 employees	Employer use of NCQA accreditation and HEDIS measures in health plan selection
Meyer et al.	1998	Case studies and focus groups	5 public and private health plan purchasing initiatives	Consumer use and attitudes about publicly available comparative health plan information
Chernew and Scanlon	1998	Conditional logit and market- share regression analysis	National sample of employees choosing single MC coverage in 1995 through a Fortune 50 company	Impact of health plan ratings on MC plan enrollment
Scanlon and Hendrix	1998	Survey	20 large purchasers of employee health benefits	Use of accreditation and performance measures in purchasing decisions
Hibbard et al.	1998	Telephone survey	Medicare beneficiaries	Knowledge of difference between fee-for-service and MC plans
Robinson and Brodie	1997	Survey	Health care consumers	Importance and use of quality information
Tumlinson et al.	1997	Survey	Massachusetts state employees	Views on plan performance information
Gibbs et al.	1996	Focus groups	Medicare, Medicaid, and privately insured enrollees	Information needed for health plan choice
Sainfort and Booske	1996	Study of hypothetical plan choice	Wisconsin state employees	Use of computerized consumer information in plan choice
Edgman-Levitan and Cleary	1996	Literature review and interviews	Published and unpublished studies and interviews conducted with organizational representatives	Information needs of and use by consumers
Jewett and Hibbard	1996	Focus groups	Medicare, Medicaid, and privately enrolled consumers	Comprehension of quality-of- care indicators
Hibbard and Jewet	1996	Focus groups and survey	Privately insured, uninsured, and Medicaid recipients in Oregon	Usefulness and salience of quality measures used in plan report cards
McGee and Knutson	1994	Literature review	Published and unpublished studies	Consumer view of health plan report cards

regarding the value of expert opinions and the opinions of peers obtained through consumer surveys. The findings on the salience of advice from family and friends are particularly important for revealed preference studies because important “informal information” is treated as a randomly distributed unobserved variable in these studies.

Several studies have asked consumers about the types of information desired or the ability to understand commonly reported measures. Gibbs et al. (1996) presented focus group members with two examples of health plan report cards. One report card contained several categories of aggregated measures, while the other contained more detailed data on the disaggregated measures. Those who said that they would use the report cards when choosing a health plan preferred the more detailed, disaggregated report card. Several studies found that respondents were skeptical about a plan’s ability to influence many of the measures contained in health plan report cards. For example, respondents believe that utilization of preventive care services, such as immunizations, is the responsibility of the enrollee rather than the plan. Some authors have suggested that respondents’ failure to see a plan’s responsibility for such measures is likely due to a poor understanding of MC and health insurance in general. In a study of Medicare beneficiaries, Hibbard and Jewett (1998) report that only 11% of respondents had knowledge adequate to make an informed choice between traditional Medicare coverage and a Medicare risk plan. Jewett and Hibbard (1996) found that focus group participants from three groups—Medicare, Medicaid, and privately insured—had such a poor understanding of health plan quality measures that it was unlikely they were capable of using comparative report cards appropriately.

Two studies (Sainfort and Booske 1996; Booske et al. 1999) examined the hypothetical enrollment choices made by 202 state employees in Wisconsin who used a computerized plan education system. Although their choices were nonbinding, the choice offerings included actual plans available to state employees during the open enrollment period, so it is likely that the hypothetical decision-making process was salient and meaningful.

The computerized decision tool provided participants with the opportunity to obtain as little or as much comparative plan information as desired. For each respondent, the authors recorded the type of information accessed, the amount of time spent examining information, and the final enrollment decision. Analysis of the data indicated that the most important

plan attributes were benefits/coverage, costs, and provider participation. The authors also found clear relationships between plan choice and the amount of information accessed and time spent analyzing the information. Sainfort and Booske concluded that consumers appear to use information if they examine comparative reports. Hence, policy efforts aimed at increasing competition in health insurance markets must not just provide information to consumers, but must also ensure that the provided information is understood and used.

A study by Knutson et al. (1998) surveyed two groups of Minnesota state employees to ascertain the impact of a survey-based report card. One group, experimental, received the report card while the other group, controls, did not. The authors compared the responses of the experimental with those of the controls using a series of statistical tests and regression analyses. The authors concluded that “the report card had few discernible effects on employees’ knowledge, attitudes, or choice of health plans” (p. 23).

A laboratory experiment examining the effect of a CAHPS-based report yielded some positive results. Spranca et al. (2000) randomly assigned a sample of 311 adults between the ages of 18 and 64, who were recruited to participate in the experiment, to receive various forms of enrollment materials for hypothetical health plans before being asked to make a nonbinding plan choice. Participants were presented with different choice sets that varied by plan type, covered benefits, out-of-pocket costs, and provider ability. The experimental groups received CAHPS reports in addition to basic enrollment materials. The authors report that the CAHPS ratings influenced plan choice when low-cost plans received high CAHPS ratings. Interestingly, the CAHPS report did not affect plan choice for more expensive plans. The authors interpret these results as evidence that plan ratings can facilitate cost containment by moving consumers into low cost-high quality plans.

Accurately projecting the impact of performance measures on plan enrollment and financial stability requires an understanding of the characteristics of those who are most likely to use health plan report cards and performance measures. A 1998 report by the NCQA on consumer use of report cards sheds light on this issue. As part of the Consumer Information Project, the NCQA conducted focus groups to determine what information consumers desired about health plans. Based on these responses, the NCQA designed plan report cards for two markets—Denver and St. Louis—and distributed the reports through pur-

chasers and purchasing coalitions. The NCQA then surveyed consumers from both markets to determine if and how the report cards were used.

Fifty percent of respondents in the combined sample (both Denver and St. Louis) remember receiving the report cards. Of this group, 40% indicated that they read most or all of the report card and 66% indicated that the reports were helpful in deciding whether to stay in their current plan or switch plans. The probability of reading the comparative reports was not influenced by age, education, number of dependents, self-reported health status, presence of chronic health conditions, or past utilization.

Surveys were administered to respondents in Denver before and after distribution of the report cards to test whether perceptions of health plan quality and performance measurement changed as a result of seeing the comparative plan reports. Three measures were significantly more important in the post-report card survey relative to the pre-report card survey: NCQA accreditation status, plan disenrollment due to dissatisfaction, and childhood immunization rates. A St. Louis survey asked respondents if they had used the report card when making enrollment decisions. In general, the report cards did not have as much of an effect on plan enrollment as anticipated. There were no differences between respondents who indicated using the report cards when making enrollment decisions and those who did not with regard to sex, self-reported health status, presence of chronic health conditions, or health care utilization. However, differences did exist for age (middle-aged respondents were more likely than older or younger respondents to use the report card), coverage category chosen (those choosing family coverage were more likely to use the report card), and tendency to use consumer reports for other major purchasing decisions (those who frequently use consumer reports for other products and services were more likely to use the report card).

Some analysts believe that the true audience for comparative health plan information is not the end consumer but those who purchase health insurance on behalf of the end consumer (for example, corporations, Medicare, and state Medicaid programs). Two studies provide evidence on the usefulness of health plan accreditation, HEDIS, and performance measurement in the purchasing community. In a random survey of 1,502 public and private employers conducted in 1997, Gabel et al. (1998) found little evidence that health plan accreditation status or HEDIS measures were important in deciding which health plans to offer employees. Only 11% of respondents indicated that

NCQA status was very important in the decision making process, and 5% indicated the HEDIS results were very important. One percent of the surveyed firms provided HEDIS data to employees to assist in plan selection.

In another survey of purchasers, Scanlon and Hendrix (1998) found stronger evidence that HEDIS results and NCQA accreditation status influenced contracting decisions. However, the sample in their study was smaller and more representative of very large purchasers than the Gabel et al. study. The authors suggest caution in interpreting the results because the respondents included in their study are known to be actively involved with HEDIS and accreditation initiatives, likely biasing the results for the importance of performance measurement upwards.

Findings from Revealed Preference Studies

Only three studies empirically examined the relationship between actual binding health plan choices and plan performance measures or ratings. Farley et al. (1999) report preliminary results of a CAHPS demonstration and evaluation project in New Jersey where new Medicaid beneficiaries were randomly assigned to control and experimental groups, with the experimental group receiving health plan ratings for the relevant Medicaid plans. The plan ratings were based on the survey responses of Medicaid plan enrollees in New Jersey. The control group did not receive the CAHPS-based ratings. Both the control and experimental groups were asked to select an MC plan; otherwise they were auto-assigned to a plan by the state. Farley and her colleagues found no evidence that the CAHPS-based ratings influenced plan choice when the entire sample was analyzed: neither the auto-assignment rate nor enrollment choices differed significantly between the experimental and control groups. However, when only a subset of the sample that remembered seeing and using the report was analyzed (identified through a follow-up survey), a CAHPS effect was found. For example, this “exposed” group had a lower rate of auto-assignment and was more likely to choose health maintenance organizations with higher CAHPS ratings.

Chernew and Scanlon (1998) examined the relationship between MC enrollment and employer-provided health plan ratings for single employees of a Fortune 50 firm. The ratings were constructed by the firm using HEDIS data and were provided to employees during the annual open enrollment period. The authors estimated conditional logit and share-regression models to determine whether the probabil-

ity of plan enrollment was influenced by the five plan ratings (enrollee satisfaction, surgical quality, medical treatment, physician quality, and preventive care), out-of-pocket price, model type, and availability of physicians. The models were specified such that the probability of enrollment for any given employee in any given plan was not only a function of the plan's characteristics (price and performance), but also of the characteristics of other plans offered to the employee. Because some employees have the same plan choices, models that use individual level data will generate a correlation in unobserved variables across observations, resulting in downward-biased standard errors. Aggregation of the individual data to "market shares" eliminates this problem.

The authors estimated an inverse relationship between out-of-pocket price and plan enrollment, but did not find the hypothesized positive relationship between ratings and enrollment. Paradoxically, the authors found several counterintuitive but statistically significant relationships between plan enrollment and plan ratings, suggesting the importance of unobserved plan attributes in the decision-making process. For example, the surgical care rating, which measured utilization of surgeries thought to be potentially overused, and the satisfaction rating were both estimated as inversely related to the probability of plan enrollment.

In a more detailed study, Scanlon and Chernew (1999) extended their earlier analysis using 1996 (as opposed to 1995) enrollment and rating data and the associated underlying HEDIS data. The more recent ratings were better for three reasons. First, the underlying HEDIS data were more complete with regard to specific measures and for entire plans. Second, the 1996 ratings indicated performance on a three-interval scale for each plan in each domain of performance. In 1995 only two categories of performance existed for each domain, and plans were identified only if they were in the better category, "superior." Hence, failure to achieve the superior designation could only be inferred by omission of the indication of superior. Third, by 1996 employees were more familiar with the rating and reporting process.

The availability of the underlying HEDIS measures in addition to the ratings allowed examination of hypotheses generated from the earlier work that could not otherwise be explored. For example, the inverse relationship between enrollment and the surgical care rating may have reflected the fact that the indices constructed by the firm penalized plans performing more surgeries while rewarding plans performing

fewer. If consumers observe and value easy access to specialists, and specialist access is positively correlated with surgical volume, then one might expect to observe a counterintuitive sign. Similarly, if employees prefer popular physicians, and popular physicians have longer waiting times or do not see new patients (two measures used to construct the satisfaction rating), one might expect an inverse relationship between the satisfaction rating and enrollment.

The results from the extended work found no evidence that employees relied heavily on report cards, despite the better presentation and greater familiarity with the performance measurement process. Estimates using just the newer, more detailed ratings failed to identify any consistent positive relationship between better performance and probability of enrollment. Only the satisfaction domain demonstrated the hypothesized relationship, and the authors attribute this switch in sign from the earlier work to changes in rating construction (discussed below). The surgical rating continued to display counterintuitive signs, and the other domains showed no strong relationship to enrollment.

The analysis of the relationship between the underlying HEDIS measures and enrollment confirmed earlier hypotheses. For example, analysis of the underlying HEDIS data used to construct the surgical rating indicated that the inverse relationship between the surgical rating and enrollment was not a response to any particular type of surgery, but a broad-based correlation between higher surgical rates and probability of enrollment. This supports the hypothesis that surgery rates act as proxy for access to specialists, and employees appear to prefer easy access to specialists.

Similarly, the analysis of detailed HEDIS data indicates a positive, though generally not statistically significant, relationship between satisfaction measures and enrollment. Variables included in the earlier satisfaction rating, such as percentage of physicians accepting new patients, were inversely correlated with the probability of enrollment, confirming our hypothesis that physician popularity was an important determinant of plan choice and inversely correlated with the HEDIS measure recording the percentage of a plan's physicians accepting new patients.

A third type of analysis included both ratings and the underlying HEDIS measures. Presumably the underlying HEDIS measures would capture some of the unobserved plan traits with which they were correlated, leaving the rating variables to capture the impact of information about plan performance. This analysis did not strengthen the results. Employees still

appeared not to be strongly influenced by health plan performance ratings.

5. CONCLUSION

The results of this study indicate that although health plan report cards and performance measures are being widely disseminated to promote competition among plans in health insurance markets, there is little current empirical evidence that consumers use this type of information in plan selection. In fact, most evidence indicating that consumers may use plan performance information comes from focus group or survey data in which the sample size is so small that it is difficult to interpret the findings statistically rather than anecdotally.

Moreover, those studies suggesting that consumers do use plan performance information do not find significant differences between those who do and do not use this information with regard to important demographic characteristics used by actuaries, such as age, sex, health status, and prior health-care utilization. Thus, although performance measures have the potential to influence plan enrollment and therefore the demographic composition of a plan's enrollees, there is currently no evidence that plans scoring "good" or "bad" systematically receive one type of enrollee (for example, sicker patients).

Despite the lack of evidence that consumers respond to plan ratings, it would be premature to conclude that plan performance ratings are useless or can never affect plan enrollments, for three reasons. First, most of the current empirical evidence has evaluated the very earliest attempts to provide consumers with report cards. As reporting and dissemination efforts are refined and consumers become more familiar with comparative plan information, more of an impact on plan enrollments could be detected. Second, most of the existing studies analyze data obtained from focus groups, surveys, or laboratory experiments rather than binding enrollment decisions. Even the studies that have used revealed preference data are cross-sectional studies and therefore may be biased by lack of control for important unobserved variables, such as the opinions of family members or physician participation in a given plan. Third, having to report measures such as HEDIS and having plan performance measures publicly disseminated in the form of report cards could yield benefits other than increased enrollment. For example, if performance measures help plans identify problem areas and ultimately lead to improved patient

care or cost savings, plans, purchasers, and consumers may benefit from performance initiatives.

In short, researchers should continue to monitor the impact of performance measures for managed health-care plans because the political and economic climate suggests that initiatives to collect, report, monitor, and disseminate plan performance information will likely continue, and it is possible that consumers will begin to find this information useful.

ACKNOWLEDGMENTS

We are grateful to the SOA for funding this research. We acknowledge the helpful comments of Tom Edwards and members of the HEDIS Project Oversight Group. We also appreciate the comments of participants at the "Managed Care in a Time of Transition" meeting in Minneapolis, May 4–5, 1998 and at the SOA annual meeting in New York, October 1998.

REFERENCES

- BOOSKE, B.C., SAINFORT, F., AND HUNDT, A.S. 1999. "Eliciting Consumer Preferences for Health Plans," *Health Services Research* 34, no. 4 (October):839–54.
- CHERNEW, M.E., AND SCANLON, D.P. 1998. "Health Plan Report Cards and Insurance Choice," *Inquiry* 35 (spring):9–22.
- CLEARY, P.D., AND EDGMAN-LEVITAN, S. 1997. "Health Care Quality: Incorporating Consumer Perspectives," *Journal of the American Medical Association* 278(19):1608–11.
- DONABEDIAN, A. 1988. "The Quality of Care: How Can It Be Assessed?" *Journal of the American Medical Association* 260, no. 2 (September 23–30):1743–48.
- DONABEDIAN, A. 1989. "The End Results of Health Care: Ernest Codman's Contribution to Quality Assessment and Beyond," *Milbank Quarterly* 87:233–55.
- EDGMAN-LEVITAN, S., AND CLEARY, P.D. 1996. "What Information Do Consumers Want and Need?" *Health Affairs* 15, no. 4: 42–56.
- ENTHOVEN, A.C. 1988. "Theory and Practice of Managed Competition in Health Care Finance." Professor Dr. F. de Vries Lectures. North Holland/American Elsevier.
- ENTHOVEN, A.C. 1989. "Management of Competition in the FEHBP," *Health Affairs* Fall:33–50.
- FARLEY, D.O., SHORT, P.F., ELLIOT, M., KANOUSE, D., BROWN, J., AND HAYS, R.D. 1999. "Use of CAHPS Information in Health Plan Choices By New Jersey Medicaid Beneficiaries." Unpublished Manuscript (draft). Santa Monica, Ca: RAND Corporation.
- GABEL, J.R., HUNT, K.A., AND HURST, K. 1998. "When Employers Choose Health Plans: Do NCQA Accreditation and HEDIS Data Count?" New York: Commonwealth Fund Report.
- GIBBS, D.A., SANGL, J.A., AND BURRUS, B. 1996. "Consumer Perspectives on Information Needs for Health Plan Choice," *Health Care Financing Review* 18(1):55–73.

- HIBBARD, J.H., AND JEWETT, J.J. 1996. "What Type of Quality Information Do Consumers Want in a Health Care Report Card?" *Medical Care Research and Review* 53:28-47.
- HIBBARD, J.H., JEWETT, J.J., ENGELMANN, S., AND TUSLER, M. 1998. "Can Medicare Beneficiaries Make Informed Choices?" *Health Affairs* 17(6):181-93.
- IEZZONI, L.I. 1996. "100 Apples Divided by 15 Red Herrings. A Cautionary tale from the Mid-19th Century on Comparing Hospital Mortality Rates," *Annals of Internal Medicine* 124(12):1079-85.
- JEWETT, J.J., AND HIBBARD, J.H. 1996. "Comprehension of Quality Indicators: Differences Among Privately Insured, Publicly Insured, and Uninsured," *Health Care Financing Review* 18(1):75-94.
- KNUTSON, D.J., KIND, E.A., FOWLES, J.B., AND ADLIS, S.A. 1998. "Impact of Report Cards on Employees: A Natural Experiment," *Health Care Financing Review* 20(1):5-27.
- KPMG PEAT MARWICK LLP. 1998. *Health Benefits in 1998: Executive Summary*. Arlington, Va.:KPMG.
- MAINOUS, A.G., AND TALBERT, J. 1998. "Assessing Quality of Care via HEDIS 3.0: Is There a Better Way?" *Archives of Family Medicine* 77:410-13.
- MCDONALD, L. 1998. "Florence Nightingale: Passionate Statistician," *Journal of Holistic Nursing* 16(2):267-77.
- MCGEE, J., AND KNUTSON, D. 1994. "Health Care Report Cards: What About Consumers' Perspectives?" *Journal of Ambulatory Care Management* 17(4):1-14.
- MCGLYNN, E.A. 1998. "Choosing and Evaluating Clinical Performance Measures," *Joint Commission Journal on Quality Improvement* 24(9):470-79.
- MGLENDON, W.W. 1990. "Ernest A. Codman, MD (1869-1940), the End Result Idea, and *The Product of a Hospital*," *Archives of Pathology and Laboratory Medicine* 114 (November):1101-4.
- MEYER, J.A., WICKS, E.K., RYBOWSKI, L.S., AND PERRY, M.J. 1998. "Report on Report Cards: Volume I." Washington, D.C.: Economic and Social Research Institute.
- NATIONAL COMMITTEE FOR QUALITY ASSURANCE. 1997. "HEDIS 3.0." Washington, D.C.: NCQA.
- NATIONAL COMMITTEE FOR QUALITY ASSURANCE. 1998. "Do Consumers Use Health Plan Report Cards?" Washington, D.C.: NCQA.
- NATIONAL COMMITTEE FOR QUALITY ASSURANCE. 1999. <http://www.ncca.org/pages/policy/accreditation/mco/faq.htm>.
- PYENSON, B.S., DUBAUSKAS, M.A., YAMAMOTO, D.H., TWISS, A.K., AND ALLEN, J.G. 1998. "The Actuary's Role in Managed Care," *North American Actuarial Journal* 2, no. 3 (July): 128-36.
- ROBINSON, S., AND BRODIE, M. 1997. "Understanding the Quality Challenge for Health Consumers: The Kaiser/AHCPR Survey," *Joint Commission Journal on Quality Improvement* 23:239-44.
- SAINFORT, F., AND BOOSKE, B.C. 1996. "Role of Information in Consumer Selection of Health Plans," *Health Care Financing Review* 18(1):31-54.
- SCANLON, D.P., AND CHERNEW, M.E. 1999. "HEDIS Measures and Managed Care Enrollment," *Medical Care Research and Review* 56 (Supplement 2):60-84.
- SCANLON, D.P., AND HENDRIX, T.J. 1998. "Health Plan Accreditation: NCQA, JCAHO, or Both?" *Managed Care Quarterly* 6(4):52-61.
- SCANLON, D.P., CHERNEW, M.E., SHEFFLER, S., AND FENDRICK, A.M. 1998. "Health Plan Report Cards: Exploring Differences in Plan Ratings," *Joint Commission Journal on Quality Improvement* 24 (January):5-20.
- SHORT, P.F. 1998. "Early Lessons from the CAHPS Demonstrations and Evaluations," CAHPS Users Meeting sponsored by the Agency for Health Care Policy and Research, Baltimore, Md.
- SPRANCA, M., KANOUSE, D.E., ELLIOT, M., SHORT, P.F., FARLEY, D.O., AND HAYS, R.D. 2000. "Do Consumer Reports of Health Plan Quality Affect Health Plan Selection?" *Health Services Research* (forthcoming).
- TUMLINSON, A., BOTTIGHEIMER, H., MAHONEY, P., STONE, E.M., AND HENDRICKS, A. 1997. "Choosing a Health Plan: What Information Will Consumers Use?" *Health Affairs* 16:229-39.
- WICKS, E.K., MEYER, J.A., RYBOWSKI, L.S., AND PERRY, M.J. 1999. *Report on Report Cards: Volume II*. Washington, D.C.: Economic and Social Research Institute.

Discussions on this paper can be submitted until October 1, 2000. The authors reserve the right to reply to any discussion. See the Submission Guidelines for Authors on the inside back cover for detailed instructions on the submission of discussions.