California Provider Group Report Cards: What Do They Tell Us?

Lisa Payne Simon, MPH, and Ann F. Monroe, MA

The objective of this study was to perform a practical assessment of publicly reported data from 4 reports on California provider groups through the eyes of the consumer. The study compared performance indicator content and rating methodologies, examined the degree of correlation in provider group performance on indicators common to 2 or more reports, and assessed the level of concordance among summary ratings of performance. Comparative analyses revealed significant variation in performance indicator content, data sources, and rating methodologies. Spearman correlation analysis revealed highly correlated group performance on patient satisfaction and member-requested group transfers, poorly correlated performance on breast and cervical cancer screening, and moderately correlated performance on state and regional average scores. Summary ratings applied to these data were only moderately correlated. These findings suggest that competing California provider group report cards produce inconsistent messages about provider quality and may create barriers to use, comprehension, and reliance upon quality information among consumers and other potential users.

Key words: Health maintenance organization, Provider group, Quality, Report card, Variation.

The measurement of health care quality performance is a relatively new science. Like many new scientific endeavors, the field of performance measurement is both full of promise and fraught with complexities. In recent years, a number of popular journalism efforts at quality performance measurement have surfaced, such as those from *US News & World Report* (1) and *Consumer Reports* (2). These efforts represent pioneering attempts to report health care quality performance information to the public and to promote quality accountability among health plans and providers.

Increasing emphasis on accountability and a growing national interest among consumers and other stakeholders for information on physician quality performance have led to the proliferation of provider profiling attempts. Physician profiling is now widely used by many health care systems in the United States. Accompanying this trend is a growing debate over the adequacy and appropriateness of provider profiling. Fueling this debate is the fact that very little is known about the reliability of commonly used profiling systems. For example, Hofer et al (3) recently documented the unreliability of individual physician "report cards" for assessing the costs and quality of diabetes care. In a cohort study performed in 3 diverse practice settings, Hofer found that physician report cards for diabetes care were unable to detect reliably true practice differences within the 3 sites studied. The growing controversy around physician profiling is exemplified by the author’s conclusion that use of individual physician profiles may foster an environment in which physicians can most easily avoid being penalized by avoiding or deselecting patients with high prior cost, poor adherence, or poor response to treatment.

In addition to growing controversy about provider report cards, there is, at the same time, growing competition over who should measure provider quality. Over the past few years, the accreditation marketplace has been characterized by commercial struggle over the franchise of provider measurement among 3 leading accreditation agencies: the American Medical Association (AMA), the Joint Commission on Accreditation of Healthcare Organizations (JCAHO), and the National Committee for Quality Assurance (NCQA). Just recently, after investing 4 years and $12 million, the AMA discontinued its American Medical Accreditation Program (AMAP). At the top of the list of reasons for the demise of AMAP were finances, politics, and the volatility of the measurement and accreditation marketplace (4).
Because of some relatively unique characteristics of the California managed care marketplace, measurement of the quality of providers in California is typically applied at the level of the provider organization. The California provider group market is characterized by increasing capitation under managed care, mergers and consolidations into large provider groups, and network model contracting. Through network model contracting, a single provider group contracts with multiple, competing health plans. Other US markets have witnessed some of these changes but have not witnessed all of these changes or witnessed the changes to the same degree as California has.

Largely as a result of these developments, attention to quality at the provider group level has dramatically increased in recent years, particularly in California. For example, in 1996, accrediting bodies such as NCQA and JCAHO began offering provider organization certification and network accreditation programs. Also in 1996, the Pacific Business Group on Health (PBGH) implemented data collection among provider groups on the West Coast using a standardized patient (PBGH) implemented data collection among provider groups on the West Coast using a standardized patient experience survey, the Physician Value Check (5).

Some health plans have established proprietary provider profiling systems as a service to their enrollees and to differentiate themselves in the marketplace. In California, 3 large commercial health plans annually collect and publicly report quality performance data on their provider networks that significantly overlap. As a result, many of the same provider groups are included in each of these reports. Data included in the California health plan provider network reports are derived from member surveys and other plan-specific sources.

Many California consumers choose a new health plan, a new health provider group, or both during annual open enrollment season or whenever they switch jobs. At these times, many consumers receive from their employer report card information on multiple plans and provider networks from which they can choose. Using this information, consumers can compare and contrast plan and provider network quality performance and may wish to include quality performance information in their selection process.

The following scenario is intended to illustrate how California consumers might actually use plan and provider network report card information. New employees must choose a health plan and provider group for themselves and their families. The employee can choose from among 3 health plans—Blue Cross of California, Health Net, and PacifiCare—and from a large number of provider groups located in the employee’s metropolitan area. The employer, a large corporation, provides the employee with quality performance information produced by each of the 3 health plans and by PBGH, including a list of 1999 PBGH Blue Ribbon Award–winning provider groups. Publicized each year at health plan open enrollment period, the PBGH Blue Ribbon Award is widely regarded as a “gold standard” among California purchasers, consumers, and health system providers.

Developed from actual California report card data and in collaboration with PBGH, Table 1 illustrates the performance of PBGH Blue Ribbon provider groups by other California ranking systems. Table 1 demonstrates the dilemma faced by our consumer who is attempting to select a provider group based on publicly reported health plan materials: quality performance results are inconsistent for 5 of the 6 groups listed in Table 1. Given the observed lack of concordance and reliability among these reports, what are the implications for informed consumer choice?

In this study, we performed a practical assessment of publicly reported data from 4 reports on California provider groups: Blue Cross of California’s Provider Performance Profile (6), Health Net’s Participating Physician Group Report Card (7), PacifiCare’s Quality Index (8), and PBGH’s Physician Value Check (5). Ex-

---

**Table 1**

Performance of PBGH Blue Ribbon Provider Groups as Assessed by Other California Ranking Systems*

<table>
<thead>
<tr>
<th>PBGH Blue Ribbon Provider Group</th>
<th>Blue Cross of California (Excellent/Good/ Satisfactory)</th>
<th>Health Net (Excellent/Very Good/ Good)</th>
<th>PacifiCare of California (Percentile; Total No. Best Practices)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alta Bates Medical Group</td>
<td>Good</td>
<td>Not included</td>
<td>67th percentile; no Best Practices</td>
</tr>
<tr>
<td>Bristol Park Medical Group</td>
<td>Good</td>
<td>Excellent</td>
<td>64th percentile; 1 Best Practice</td>
</tr>
<tr>
<td>Brown &amp; Toland Medical Group</td>
<td>Excellent</td>
<td>Excellent</td>
<td>57th percentile; 1 Best Practice</td>
</tr>
<tr>
<td>Hill Physicians Medical Group</td>
<td>Excellent</td>
<td>Excellent</td>
<td>42nd percentile; no Best Practices</td>
</tr>
<tr>
<td>Huntington Provider Group</td>
<td>Good</td>
<td>Good</td>
<td>15th percentile; no Best Practices</td>
</tr>
<tr>
<td>Palo Alto Medical Foundation</td>
<td>Excellent</td>
<td>Excellent</td>
<td>Not included</td>
</tr>
</tbody>
</table>

*PBGH indicates Pacific Business Group on Health. Terms in parentheses indicate possible ratings.
examining these data through the eyes of the consumer, we compared performance indicator content and rating methodologies; we examined the degree of correlation in provider group performance on indicators common to 2 or more reports; and we assessed the level of concordance among summary ratings of performance. Factors contributing to the observed lack of concordance between these reports were identified. We conclude with a discussion of the implications of competing performance reports, and recommendations for improved measurement and public accountability efforts.

METHODS

California has 4 separate publicly reported provider group quality performance reports. A brief description of each of these reports is summarized below.

- In 1999, Blue Cross of California published Provider Performance Profile—a statewide comparative report on the quality of 134 provider groups in the Blue Cross network.
- Health Net’s Participating Physician Group Report Card is a regional comparative report of members’ satisfaction ratings of provider groups included in the network. There are 131 provider groups included in Health Net’s most recent public report released in 1998.
- PacifiCare of California’s Quality Index is a comparative statewide report on the quality of provider groups in the PacifiCare network. Included in the 1999 public report are 187 contracting provider groups. PacifiCare has released the Quality Index annually since 1998.
- The PBGH’s Physician Value Check is a regional comparative report of provider group performance on patients’ satisfaction with care and assessment of health services delivered. The PBGH has conducted the Physician Value Check survey semiannually since 1996. Included in the 1999 report are 51 provider groups from California and 7 from the Pacific Northwest.

In this study, we gathered the most recent public provider group reports for comparative analysis. We identified 6 indicators that—at face value—appeared comparable and were common to 2 or more reports: patient satisfaction with the medical group, member-requested group transfers, cervical cancer screening, breast cancer screening, regional overall performance, and California overall performance.

Results from each report were compared at the provider group level. Because of the size and geographic spread of the health plan networks, there were a sufficient number of overlapping groups for comparative analysis. For example, the Blue Cross of California Provider Performance Profile and the PacifiCare Quality Index had 103 provider groups in common. Information on provider report card methods (including sample size, sampling frames, treatment of missing data, risk adjustment methods, scoring, and statistical methods) were obtained from public reports and/or through personal communications with report card developers.

We compared provider group performance among groups included in more than 1 report card and across 3 categories of publicly reported data as follows:

1. Performance on 4 individual indicators: satisfaction with the medical group, member-requested group transfers, breast cancer screening, and cervical cancer screening;
2. Performance on aggregate measures: regional and statewide average scores, and;
3. Performance on summary ratings: qualitative designations applied to numerical performance scores, for example, “Above Average,” “Below Average,” “Best Practice,” and “Satisfactory.”

Spearman correlation analyses were performed to evaluate the degree of correlation in provider groups’ rankings on individual and aggregate performance indicators. All statistical analyses were performed with STATA statistical software (Version 5.0) (9). A comparative assessment of the level of concordance among provider groups’ summary ratings was also performed. Lastly, a series of case studies were performed from the perspective of the consumer who might use these data to evaluate or compare provider group performance.

RESULTS

Comparative analysis of California provider group report card content and methodology revealed significant variation. Table 2 illustrates the variety of indicators and rating methodologies observed in each of the 4 reports.

Table 3 demonstrates variation in definitions and data sources among 6 indicators common to 2 or more reports; it shows that most of these “common” indicators were not consistent with respect to definitions, data collection, or statistical methods employed. However, these 6 measures might appear comparable to the average user. For example, the Blue Cross of California, PacifiCare, and PBGH provider group reports each include a breast cancer screening performance in-
Table 2
California Provider Group Report Card Contents and Rating Methodologies

<table>
<thead>
<tr>
<th>Reporting Initiative</th>
<th>Quality Performance Indicators</th>
<th>Summary Ratings</th>
<th>Summary Rating Methodology</th>
</tr>
</thead>
</table>
| Blue Cross of California Provider Performance Profile | ● Provider qualifications  
● Grievance and appeal overturned  
● Grievance turnaround time  
● Member retention (member-requested group transfers)  
● Access and service (patient satisfaction with group)  
● Staying healthy: Smoking inquiry  
Breast cancer screening  
Cervical cancer screening | ● Overall regional performance  
Overall performance in California | Three performance categories:  
Excellent (score >80 out of 100 total points)  
Good (score 20–80 out of 100 total points)  
Satisfactory (score <20 out of 100 total points) |
| Health Net Participating Physician Group Report Card | ● Member satisfaction with:  
Quality of care  
Access to care  
Medical group | Regional rating (Bay Area, Central Valley, Los Angeles, San Diego, Sacramento) | In regional comparisons, tests of statistical significance used to determine 3 performance categories:  
Excellent  
Very Good  
Good |
| PacifiCare of California Quality Index | ● Member satisfaction with:  
Medical group  
Primary care physician  
Timeliness of referral process  
● Member complaints  
● Medical group transfers  
● Overturned appeals  
● Breast cancer screening  
● Cervical cancer screening  
● Eye exam for diabetic patients  
● Optimal outpatient care to avoid the need for hospitalization  
● Prescription drug treatment for congestive heart failure  
● Encounter data | ● Overall service score  
Overall clinical score  
Overall statewide score | Percentile rank based on the average of 14 performance indicator scores.  
Percentile ranks equal to or greater than 90 are designated “Best Practice” performers. |
| Pacific Business Group on Health Physician Value Check | ● Overall satisfaction  
● Overall rating of care  
● Ease of getting care  
● Promptness of care  
● Quality of care  
● Doctor’s communication skills  
● Courtesy of office staff  
● Cervical cancer screening  
● Breast cancer screening  
● Colon cancer screening  
● Explanation of hormone replacement therapy  
● Flu shots for seniors | No summary ratings applied | In regional comparisons, tests of statistical significance used to determine 3 performance categories:  
Above Average  
Average  
Below Average |

dicator. PacifiCare’s indicator reports performance based on the NCQA Health Plan Employer Data and Information Set (HEDIS) 1999 Breast Cancer Screening administrative data specification (ie, based on plan encounter data, the percentage of women aged 52 through 69 years who were continuously enrolled during the reporting year and the preceding year, and who had a mammogram during the reporting year or the preceding year) (10). Although the PBGH and Blue Cross of California breast cancer indicators are both derived from patient surveys, each used a different survey instrument and had different sampling frames. PBGH’s measure is derived from a representative sample of all patients of a provider group, whereas Blue Cross of California’s sample represents only their plan members assigned to a provider group. This example illustrates the range of methodologies used to generate a “common” measure included in several California provider group reports. Results of Spearman correlation analysis are sum-
Table 3
Comparison of Performance Indicators Included in 2 or More California Provider Group Reports

<table>
<thead>
<tr>
<th>Performance Measure</th>
<th>Blue Cross of California</th>
<th>Health Net</th>
<th>PaciCare</th>
<th>PBGH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Definition</td>
<td>Data Source</td>
<td>Definition</td>
<td>Data Source</td>
</tr>
<tr>
<td>Patient satisfaction with medical group</td>
<td>Overall satisfaction with access and service; average of 22 survey items; 5-point scale; percentage Excellent and Very Good reported</td>
<td><em>Physician Value Check survey</em>b</td>
<td>Rating based on the average of 2 survey items; 5-point scale; percentage Excellent, Very Good, and Good reported</td>
<td>Health Net Member Satisfaction Survey</td>
</tr>
<tr>
<td>Member-requested group transfers</td>
<td>Index based on the number of member-initiated transfers over dissatisfaction with quality of care, service, or access</td>
<td>Member database</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Cervical cancer screening</td>
<td>Percentage of returned surveys that indicate up-to-date screening status</td>
<td>Blue Cross member survey (postcard response mailer)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Breast cancer screening</td>
<td>Percentage of returned surveys that indicate up-to-date screening status</td>
<td>Blue Cross member survey (postcard response mailer)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Regional overall performance</td>
<td>Aggregate performance measure: a regional roll-up of all measures</td>
<td>Member surveys, member database, provider audit tool</td>
<td>Regional roll-up of all measures</td>
<td>Health Net Member Satisfaction Survey</td>
</tr>
<tr>
<td>California overall performance</td>
<td>Aggregate performance measure: a statewide roll-up of all measures</td>
<td>Member surveys, member database, Provider Audit Tool</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

*a* PBGH indicates Pacific Business Group on Health; NA, measure not included in health plan public report; HEDIS, Health Plan Employer Data and Information Set.

*b* Sample included Blue Cross members only for those provider groups that were not participating in the PBGH Physician Value Check initiative.
Table 4
Results of Spearman Correlation Analysis

<table>
<thead>
<tr>
<th>Performance Indicator</th>
<th>Plan-to-Plan Comparison</th>
<th>Correlation Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient satisfaction with medical group</td>
<td>Blue Cross of California vs Health Net</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>Blue Cross of California vs PaciFiCare</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>Health Net vs PBGH</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td>Health Net vs PaciFiCare</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td>PaciFiCare vs PBGH</td>
<td>0.54</td>
</tr>
<tr>
<td>Member-requested group transfers</td>
<td>Blue Cross of California vs PaciFiCare</td>
<td>0.44</td>
</tr>
<tr>
<td>Cervical cancer screening</td>
<td>Blue Cross of California vs PaciFiCare</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>Health Net vs PaciFiCare</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>PaciFiCare vs PBGH</td>
<td>0.18</td>
</tr>
<tr>
<td>Breast cancer screening</td>
<td>Blue Cross of California vs PaciFiCare</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>Blue Cross of California vs PBGH</td>
<td>–0.03</td>
</tr>
<tr>
<td></td>
<td>PaciFiCare vs PBGH</td>
<td>0.16</td>
</tr>
<tr>
<td>Regional overall performance</td>
<td>Blue Cross of California vs Health Net</td>
<td>0.40</td>
</tr>
<tr>
<td>California overall performance</td>
<td>Blue Cross of California vs PaciFiCare</td>
<td>0.37</td>
</tr>
</tbody>
</table>

*a* PBGH indicates Pacific Business Group on Health.

Table 5
Case Studies: 4 Randomly Selected California Provider Groups’ Ratings on Quality Performance Reports—Patient Satisfaction With Medical Group

<table>
<thead>
<tr>
<th>Provider Group</th>
<th>Score, %</th>
<th>Summary Rating</th>
<th>Score, %</th>
<th>Summary Rating</th>
<th>Percentile</th>
<th>Summary Rating</th>
<th>Score, %</th>
<th>Summary Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>93</td>
<td>Excellent</td>
<td>89</td>
<td>Very good</td>
<td>77th</td>
<td>Not Best Practice</td>
<td>82</td>
<td>Average</td>
</tr>
<tr>
<td>B</td>
<td>94</td>
<td>Excellent</td>
<td>90</td>
<td>Excellent</td>
<td>72nd</td>
<td>Not Best Practice</td>
<td>82</td>
<td>Average</td>
</tr>
<tr>
<td>C</td>
<td>83</td>
<td>Good</td>
<td>85</td>
<td>Excellent</td>
<td>8th</td>
<td>Not Best Practice</td>
<td>81</td>
<td>Average</td>
</tr>
<tr>
<td>D</td>
<td>83</td>
<td>Good</td>
<td>88</td>
<td>Very good</td>
<td>31st</td>
<td>Not Best Practice</td>
<td>84</td>
<td>Average</td>
</tr>
</tbody>
</table>

*a* PBGH indicates Pacific Business Group on Health.

Summary ratings applied to these data were moderately correlated. For example, Blue Cross of California and Health Net ratings of overall regional performance showed approximately 40% concordance. Agreement tended to be greatest at the high and low extremes of performance. Several groups performed consistently well in each of these reports, whereas a few poorer performing groups had consistently below average performance.

In addition to analysis of overall trends in provider group performance data, a series of case study comparisons were performed among randomly selected groups. The purpose of the case studies was to evaluate how this data might be perceived by consumers, purchasers, providers, and others who might use these reports to conduct their own comparative assessments of provider group performance.

Table 5 illustrates 4 randomly selected California provider groups’ performance on measures of patient satisfaction, while Table 6 shows performance on measures of cervical cancer screening. Focusing on patient satisfaction with a medical group (Table 5), provider group C scored 63% using the *Physician Value Check* administered by Blue Cross of California and received a summary rating of “Good.” This same group scored 85% on Health Net’s survey and received a summary rating of “Excellent.” Group C’s performance on the PaciFiCare survey was poor (ranked eighth from the
Table 6
Case Studies: 4 Randomly Selected California Provider Groups’ Ratings on Quality Performance Reports—Cervical Cancer Screening

<table>
<thead>
<tr>
<th>Provider Group</th>
<th>Blue Cross of California</th>
<th>PacifiCare</th>
<th>PBGH*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Score, %</td>
<td>Summary Rating</td>
<td>Score, Percentile</td>
</tr>
<tr>
<td>A</td>
<td>80</td>
<td>Excellent</td>
<td>97th</td>
</tr>
<tr>
<td>B</td>
<td>87</td>
<td>Excellent</td>
<td>97th</td>
</tr>
<tr>
<td>C</td>
<td>77</td>
<td>Satisfactory</td>
<td>80th</td>
</tr>
<tr>
<td>D</td>
<td>77</td>
<td>Satisfactory</td>
<td>67th</td>
</tr>
</tbody>
</table>

*PBGH indicates Pacific Business Group on Health.

bottom). Group C’s performance on PBGH’s Physician Value Check was 81% for which group C received an “Average” summary rating. The observed lack of concordance in group C’s patient satisfaction performance was frequently noted among other provider groups not only for patient satisfaction, but for other indicators as well. On the other hand, group B’s performance on cervical cancer screening (Table 6) was consistently strong. Examples of better concordance in report card results were observed, but with less frequency than examples of poor concordance.

DISCUSSION

The results of this study demonstrate significant variation in content, methodology, and results among California’s 4 publicly reported provider group report cards. Analysis of the degree of correlation among provider group scores revealed some highly correlated indicators, some poorly correlated indicators, moderately correlated regional and statewide aggregate measures, and (where comparison was possible) only moderate agreement on summary ratings of performance. With this amount of inconsistency, what is the California consumer to conclude about provider group performance or about the reliability, credibility, and utility of quality performance information in general?

Factors Contributing to Lack of Concordance

A number of methodological factors contributed to the observed lack of concordance in California provider groups’ performance. With the data available, it was not possible to quantify the magnitude of effect that each factor contributed to the observed variation; however, it was possible to identify determinants of performance variation.

One of the factors contributing to lack of concordance is the application of different statistical methods. For example, on overall ratings of group satisfaction, the Physician Value Check distinguishes high- and low-performing groups as those with risk-adjusted scores that are statistically significantly above or below (respectively) the regional average. In contrast, PacifiCare distinguishes “Best Practice” groups as those performing at or above the 90th percentile statewide, whereas no such distinction is made for underperforming groups.

Another important methodological difference involves sampling frames. For example, PacifiCare and Health Net each sample from plan members assigned to contracted provider groups. The Physician Value Check applies a group-based sampling representative of the entire patient population at each group. Blue Cross of California employs both sampling approaches in its administration of the Physician Value Check because not all of Blue Cross’ contracted groups independently participate in the Physician Value Check. Application of different sampling frames across these initiatives introduces the likelihood that different population profiles are included in report card denominators, thereby creating the potential for systematic differences in results based on underlying differences in the profile of measured populations.

Other factors weakening the concordance among these 4 provider group reports include (a) different weighting methods applied to measures, (b) different sample sizes and response rates, (c) different methods for handling missing data, (d) differences in scale development, and (e) use of risk adjustment and methods of risk adjustment. For example, PBGH and Health Net risk adjust medical group patient satisfaction results for select demographic variables such as patient age, sex, and health status. PacifiCare and Blue Cross of California do not risk adjust their medical group patient satisfaction data. Further, in its public report, PacifiCare handles missing data by specifying “Data Below Threshold.” This designation indicates that the available volume of data is less than the volume re-
quired for reliable assessment of performance. Most other health plan provider reports do not specify methods for handling missing data in analyses. Each of these and other methodological factors lead to variation in performance ratings for the same provider groups.

A methodological limitation of the current study stems from our attempt to perform comparative analysis across measures that contained inconsistencies in content, specifications, or analytic methods. In our analyses, we were unable to present appropriate, apples-to-apples comparisons as a specific consequence of the data limitations contained in the California public reports on provider group quality performance. This limitation of the analysis reinforces our hypothesis that inconsistent public reports create barriers to use, comprehension, and reliance upon quality performance information.

Costs and Consequences of Performance Measurement and Reporting Activities

Each of the 4 report cards analyzed in this study was developed to demonstrate quality accountability. It is important that accountability efforts are recognized and rewarded by consumers and by the marketplace. However, results of this study also suggest that competing performance measurement and reporting activities have costs and consequences in the marketplace. Costs and consequences fall into 2 main categories: implications for consumers and other users who attempt to evaluate or compare provider group performance using inconsistent public sources, and administrative costs and burdens associated with redundant quality measurement and monitoring activities.

Implications for Consumers and Other Users

- Inconsistent results published for the same provider group(s) are likely to create considerable confusion for consumers and other users who attempt to compare health plan network performance, and may also reduce the perceived utility and credibility of quality performance measurement in general. The case studies illustrated in Tables 1, 5, and 6 clearly demonstrate what a bewildering experience the use of competing, inconsistent data sources might be for the average consumer attempting to choose between and among health plans and their provider networks.
- Similar demonstrations of report card inconsistency have been shown. As described in Table 1, comparisons reveal how PBGH’s 1999 Blue Ribbon Award winning groups performed on other provider group ranking systems produced by Blue Cross of California, Health Net, and PacifiCare. Results for 6 provider groups indicated that the commercial health plans’ ratings and PBGH’s ratings agreed in only 1 out of the 6 groups. In a study of open enrollment materials for Ford and General Motors, McGlynn et al (11) compared health plan performance across rating systems that applied 3 comparable measures: effectiveness of care, access to care, and consumer satisfaction. They found agreement on ratings of health plan performance for less than half of the plans. In a recent study (12), an academic medical practice in Philadelphia evaluated 5 report cards produced by 2 capitated health plans between 1994 and 1997. The authors concluded that the 5 report cards contained methodological problems that led to systematic underestimation of the group practice’s performance and that larger surveys are needed to determine the accuracy of report cards in current use. As previously discussed, Hofer et al (3) recently demonstrated the unreliability of individual physician report cards for assessing the costs and quality of care for a chronic disease.

Administrative Costs and Burdens

- Each year, California health plans spend millions of dollars collecting data on network provider group performance, analyzing the data, validating data, and producing and disseminating public reports. At the same time, 51 California provider groups have (since 1996) contributed their own funds to participate in the Physician Value Check. At this time, only Blue Cross of California recognizes network provider groups’ participation in the Physician Value Check for purposes of member satisfaction assessment. Other plans, such as Health Net and PacifiCare, administer competing, proprietary surveys across their networks regardless of whether groups voluntarily participate in the Physician Value Check.
- Another source of cost associated with the plans’ quality oversight activities involves on-site audits. Each year, most health plans in California conduct 2 or more on-site quality audits at each contracting group. At this time, there are 56 licensed health plans in California and considerable overlap among health plans’ provider group networks (13). In a recent forum conducted by the California Center for Health Improvement (14), 1 provider group reported having undergone at least 50 quality audits during 1998. Each audit takes from 1 to 3 days of provider group admin-
implement policies to rationalize plan performance measurement activities and requirements. Rationalized requirements could promote a scenario in which provider groups collect and report standardized measures of performance on clinical, service, patient experience, and other measures of relevance to provider group operations. Health plans could apply these measures toward network quality oversight and marketing requirements, instead of the competing proprietary measures in use today. Methods could also be developed for rolling up provider group level measures to the health plan level. Through this scenario, health plan resources might be freed up to lower premiums and/or to collect standardized performance data on access, customer service, and other measures of particular relevance to health plan operations. This vision has several distinct advantages over current practices: it could minimize redundant measurement expense and activity, increase the consistency of quality measures across the delivery system, and enhance the supply of information available to all stakeholders.

ACKNOWLEDGMENT

The authors are grateful to Beth McGlynn, PhD, and John Adams, PhD, of RAND; to David Hopkins, PhD, of the Pacific Business Group on Health; and to Mark Smith, MD, of the California HealthCare Foundation for their comments on earlier drafts of this manuscript. The authors thank Marvin Chow and Renee Grimani-Kardener for administrative support in the preparation of this manuscript.

References