A PRACTICAL GUIDE TO PHYSICIAN PERFORMANCE PROFILES

Vinita Bahl, DMD, MPP, and Brent C. Williams, MD, MPH

At 7:20 AM Monday Dr. Schroeder starts her week at the suburban office where she and five other general internists practice. She stops at her mailbox and on the way to her office quickly sifts through the stack of mail. An envelope addressed from QualiCare, Inc., catches her attention and she decides to take a minute to look inside.

When Dr. Schroeder joined the group practice almost a year ago, the group had just renewed its contract with QualiCare, a large health maintenance organization (HMO) in the area. She estimates that about 30% of her patients are QualiCare members. Inside the envelope is a cover letter explaining the importance of providing regular feedback to physicians regarding their practice. On the next page is a chart entitled "Clinical Performance Profile for the Period 1/01/01–6/30/01," which shows data organized according to specific HEDIS measures. Dr. Schroeder sees that the chart shows her rates on the clinical performance measures and realizes that this is the first profile she's received from QualiCare.

Dr. Schroeder reviews the chart with interest, noting the numerical data listed for each of the measures, including "Diabetic patients with ophthalmology visits during the past year" and "Eligible female patients with mammography screening within the past 2 years." For each measure, Dr. Schroeder's rate is listed next to a reference rate. A final column shows whether her rates are greater than 2 standard deviations from the reference rate. In less than a minute Dr. Schroeder recognizes that the estimates of her performance are statistically unreliable due to the small number of her patients who are covered by QualiCare. Nonetheless, the low percentages of diabetic patients with eye screening and female patients with mammography make her think. She jots down a few questions for discussion with her physician partners and administrators.

So far as her personal QualiCare profile is concerned, Dr. Schroeder wants to know how the HMO uses the report. Could it possibly be used as a basis for reimbursement rates, and if so, how can the fact that the profiles are unreliable be addressed? She also wonders what reference group QualiCare uses in its profiles. In terms of her actual practice, Dr. Schroeder wonders how she and others in the group might make immediate improvements in screening for diabetic retinopathy and breast cancer. Finally, she wants to know the true rates of diabetic screening and mammography among her patients.

Physician performance profiles are reports that describe the practice patterns of physicians using measures of resource consumption, clinical performance, or patient perceptions of care. Physician profiles originally were developed by managed care organizations (MCOs) as tools to reduce variation in utilization of health care resources across populations, which often was attributed to differences in physician practice [1]. The focus of these initial profiling efforts was cost containment, and their objective was to reduce inappropriate utilization and costs by providing feedback to physicians about their practice patterns in comparison to their peers [2]. Over the years, MCOs have used profiles to address quality issues by holding contracted physicians accountable for their performance and to inform and guide quality improvement efforts [3]. In turn, physician profiling by MCOs has spawned a growing trend for physicians to gather their own data to monitor and improve their performance [4].

In recent years, other organizations have begun producing physician profiles, including state governments, health care purchaser/payer coalitions, and for-profit Internet-based companies. These organizations recognize that most decisions about medical care are made by physicians and that data about physician performance are not adequately reflected in other types of information, such as health plan or hospital profiles. The primary goals of these profiles are to hold physicians accountable for their performance and to help guide consumers when they are choosing a physician.

As physician profiles have become more widespread, so have concerns about their reliability [1,5]. Despite these concerns, the practice of profiling is growing rapidly. In a survey from the early 1990s, more than

Vinita Bahl, DMD, MPP, Office of Clinical Affairs, University of Michigan Hospitals and Health Centers; and Brent C. Williams, MD, MPH, Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, MI.

half of U.S. physicians were the subjects of performance profiling [6]. In an environment where health care purchasers are pressuring MCOs, hospitals, and physicians to demonstrate the value of the care they deliver and where consumer interest in physician performance information is growing, physician profiling likely will become more common. Accordingly, physicians must become knowledgeable about the content and major criticisms of current profiles and understand how they, themselves, can improve the practice of profiling and use profiles to enhance patient care.

The purpose of this article is to provide a practical guide to profiling for physicians who are new to or about to enter practice. The specific objectives are to 1) highlight the content and purpose of physician profiles as well as the types of organizations that produce them, 2) describe methodologic issues surrounding the production of meaningful profiles, 3) summarize current knowledge about the effectiveness of profiles in changing physician behavior and improving clinical outcomes, and 4) provide a stepwise approach for physicians to evaluate, create, and use profiles.

What Dimensions of Care are Reflected in Profiles?

Physician profiles typically report on how an individual physician or physician group is performing in one or more of the following dimensions of care: resource consumption, clinical performance, and patient perceptions. The performance measurements often are expressed as rates, such as the proportion of a pediatrician's patient population (denominator) who received recommended immunizations (numerator) or the number of hospital days (numerator) per thousand patients in the physician's practice (denominator). Profiles usually compare an individual physician's or group's performance with the performance of a reference group, such as other physicians in the region or nation, or to an absolute standard felt to represent high-quality practice. To make fair comparisons among physicians, measures often are adjusted for the severity of illness of the patients receiving care from the target physicians, a process called *case-mix adjustment*. The format for displaying the information varies. Some organizations present the performance rates with statistical confidence intervals and detailed instructions about how they should be interpreted. Others simply emphasize the identification of poor versus excellent performance [7].

Resource Consumption

Organizations that bear a financial risk for the care delivered to members of a defined population, including most MCOs and some medical groups, often profile physicians using measures of cost and resource utilization [3,8]. As a component of cost containment efforts, organizations may use cost and utilization profiles to identify individual physicians or groups with relatively high-cost practice patterns. Such physicians may then be informed of their high-cost status to respond as they wish, or they may be given financial or other incentives to decrease costs. Examples of measures that are used to profile resource consumption are hospital days per 1000 members (ie, the physician's patients or group's MCO members), outpatient laboratory costs per member per month, and prescription drug costs per member per month. **Figure 1** shows a range of measures that might be used to profile the drug prescribing patterns of an individual physician. As a basis for comparison, an individual physician's performance is commonly shown along with a group norm.

Clinical Performance

Unlike cost and most utilization measures, clinical performance measures are used to assess whether the clinical services provided are appropriate and whether they achieve desired outcomes [9]. These measures encompass both the process and outcomes of care and are typically expressed as retrospective rates for a given period of time, as in the report received by Dr. Schroeder in the opening scenario. Process measures usually are derived from evidence-based or consensus guidelines, such as the proportion of women aged 50 years and older who received a mammogram within the past 2 years. Clinical outcome measures include mortality rates, infection rates, and so on.

The most prominent set of clinical performance measures is contained in the Health Plan Employer Data and Information Set (HEDIS), which was developed under the auspices of the National Committee for Quality Assurance (NCQA). HEDIS measures are used widely and include preventive care, process, and outcome measures for patients with chronic diseases. These measures are a primary component of the NCQA health plan accreditation program and also are used by purchasers of health care (eg, employers) to measure and compare MCO performance. When trying to improve their overall NCQA scores, MCOs often use HEDIS measures to profile individual physicians in order to improve the performance of their network of contracted physicians [3]. **Figure 2** shows how HEDIS measures might be used to inform a physician about individual patients who have not received or are due to receive specific preventive care services that are appropriate, per clinical practice guidelines, for their age or

Primary Care Physician Prescribing Profile

Sponsor:	Best Care HMO	**Number of Patients:**	8
Practice Type:	Family Practice	**Time Period:**	4/1/01–6/30/01
Prescribing Physician:	John B. Good, MD 100 First Street Anytown, USA		

	Prescribing Physician		
	This Period	**Previous Period**	**All Network Physicians**
Average cost per prescription ($)	44.58	56.08	41.65
Cost per patient per month ($)	22.29	14.02	97.00
Prescriptions per patient per month (*n*)	0.50	0.25	2.47
Formulary drugs prescribed (%)	100.00	100.00	97.78
Generic drugs prescribed (%)	25.00	0.00	50.12

	Prescribing Physician	**All Network Physicians**
Average age of patients (yr)	28	49

Figure 1. Fictitious example of a resource utilization profile, combining features of actual reports received by physicians in practice. Managed care organizations commonly review and profile the drug prescribing practices of contracted physicians within their network. In this example, data on prescriptions written by an individual physician during a 3-month period are presented in chart form, with comparisons to data from the previous 3-month period and data from all other physicians in the network. Note that this prescribing profile includes information on only 8 patients, an insufficient number to accurately reflect the physician's overall prescribing patterns.

established clinical diagnosis (eg, diabetes, asthma). Unlike the retrospective rate report received by Dr. Schroeder, which also was based on HEDIS measures, the report in Figure 2 provides prospective data that the physician could more easily act upon to improve the care received by patients in his practice.

Other prominent developers of clinical performance measures are the Joint Commission for the Accreditation of Healthcare Organizations (JCAHO) and the Foundation for Accountability (FACCT). The JCAHO "core measures" focus on inpatient care processes for several different clinical populations and will soon become a component of the hospital accreditation program. For example, one of the core measures reports the percentage of eligible heart failure patients who received a prescription for an angiotensin-converting enzyme inhibitor upon discharge from the hospital. To improve a hospital's overall score, the institution's quality improvement staff may focus on improving physician performance by reporting measurement results for each attending physician.

FACCT is a nonprofit organization whose mission is to help Americans make better health care decisions. FACCT's measurement set includes process and outcome measures for several chronic diseases. Although some of these measures overlap with HEDIS, most emphasize aspects of care that are important to consumers. For example, one FACCT measure is the proportion of asthma patients reporting 0 or 1 days lost from regular activities during the past 3 months due to asthma. FACCT developed its measurement set as a means for gathering data that state governments, purchasers, and purchaser/payer coalitions could use to help consumers select a health plan or physician.

A few physician specialty societies and organizations also are developing clinical performance measurement sets. The most noteworthy is the American Medical Association (AMA) Physician Consortium for Performance Improvement, which is working in collaboration with NCQA and JCAHO to identify, develop, and define performance measurement sets that are useful to practicing physicians. Measures for a few clinical areas are being tested, many of which overlap with HEDIS measures.

Patient Perceptions of Care

Ratings of patient satisfaction with care and service are important to health care consumers [10] and purchasers [3]. MCOs, hospitals, and purchasers all measure patient satisfaction, but MCO surveys of patient satisfaction are most relevant because they are more likely to be used to assess the performance of individual physicians. In fact, Blue Cross of California recently implemented a program to pay bonuses to individual physicians and

Health and Wellness Program (HWP)
Patient Summary Report

Sponsor:	Best Care HMO	**Practice Type:**	Family Practice
Physician:	John B. Good, MD 100 First Street Anytown, USA	**Report Date:**	06/30/01

Per 6/15/01 review of HWP database, the following patients need condition-specific services to meet clinical practice guidelines*

Patient	Service	Date Needed	Last Recorded Date of Service
SS#: DOB: 8/25/86 Phone#: Been patient since: 11/01/00	Well care visits 12-21 yr • 1 well child visit	Immediately	No record
SS#: DOB: 3/7/72 Phone#: Been patient since: 09/01/00	Cervical cancer screening • 1 Pap test	Immediately	No record
SS#: DOB: 1/26/50 Phone#: Been patient since: 11/01/97	Breast cancer screening • 1 mammogram	10/15/01	10/15/99
SS#: DOB: 4/16/45 Phone#: Been patient since: 02/01/98	Diabetes • 1 retinal exam • 1 foot exam • 2 HbA_{1c} • 1 microalbumin • 1 diabetes management plan	 8/09/01 Immediately Immediately 12/15/01 Immediately	 8/09/00 No record 09/22/99 12/15/00 No record

*Clinical practice guidelines for each measure provided in HWP Check List.

Figure 2. Fictitious example of a clinical performance feedback report that is based on HEDIS measures, showing specific patient details. Clinical performance measures (eg, HEDIS measures) often are reported as retrospective rates for all patients in a physician's practice during a given period of time. However, in this example, HEDIS measures are used to provide prospective feedback on age-appropriate or condition-specific preventive services required for individual patients to meet established clinical practice guidelines. For this report, the managed care organization would need to monitor its claims database to determine whether services that were needed during a given period of time for specific patients in this physician's practice were, in fact, provided. The report then informs the physician of the dates when the services are required to meet guidelines as well as the last recorded date of the services in their claims database. Note that the actual report would have been comprehensive for all patients in the physician's practice; a few representative patient examples are provided here for illustration purposes. Although this report is fictitious, it combines features of various actual reports to provide a sense of the type of clinical performance feedback received by physicians in practice.

physician groups with high patient satisfaction scores [11]. Currently, MCOs are using many different survey instruments, but most give patients the opportunity to rate their physician on such topics as the amount of time their doctor spends with them, how well their doctor explains things, and their ability to get an appointment within an acceptable time frame. A new survey instrument, the Consumer Assessment of Health Plans (CAHPS), was recently developed by the Agency for Healthcare Research and Quality. Like HEDIS, this survey has become a component of NCQA health plan accreditation. As a consequence, CAHPS may emerge as a standard tool for assessing member/patient satisfaction with individual physicians and physician groups.

Why Profile a Physician's Performance?

Frequently, the specific goals of physician profiles are determined by the mission of the organization that produces them and the audience for which they are intended. An MCO may produce profiles for review by physicians and administrators for use in quality improvement, to determine physician reimbursement, or to inform decisions on whether to retain physicians in its network. From the viewpoint of the physician, the purposes of profiles produced by MCOs can be seen as either low-stakes (eg, feedback for informational purposes to motivate physicians to reflect on their practice) or high-stakes (eg, as a basis for bonus payments or salary withholding). Physician profiles are also made available to audiences beyond the MCOs and the physicians who contract with them, most notably employee and employer groups to inform decisions on selection of health plans or providers. Purchaser/payer coalitions are disseminating physician profiles among the coalition members and to consumers.

One notable trend in physician profiling is the release of profiling information directly to the public. For example, some MCOs are making clinical performance profiles for groups of physicians with whom they contract publicly available. In the late 1980s and 1990s, the federal government and a few states made the first efforts to release health care performance data to the public, with an initial focus on hospital performance. In New York and Pennsylvania, these efforts moved a step further with the publication of mortality rates following coronary artery bypass grafting (CABG) procedures for individual surgeons [12,13]. During the past several years, a new approach to dissemination of profiling data has emerged with the creation of Internet-based companies that provide physician-specific performance information to consumers. This information is based on satisfaction ratings from consumers and, occasionally, on the physicians' self-reported clinical performance.

Publicly available profiles are intended to promote both consumer choice and physician accountability and focus on measures of clinical performance and/or patient satisfaction. Despite growing evidence that consumers make choices about providers based on word of mouth or personal experience rather than on "hard numbers" [10], an increasing number and variety of organizations are gathering and publicly releasing physician performance data. Several examples of organizations that make profiling information available to broader audiences are listed in **Table 1**.

It is clear that physicians' views of profiles differ depending on the organization doing the profiling and on the goals and intended audience. Profiles that generate the most controversy are those that affect a physician's reimbursement or opportunity to contract with an MCO and profiles that are published for consumers [1]. This controversy stems in part from the growing recognition of the methodologic problems associated with producing reliable profiles.

Methodologic Issues: Do Profiles Reflect Actual Practice?

When a physician is profiled, he must decide how to respond to the profile report. The appropriate response depends on the purpose of the profile (high-stakes or low-stakes) and whether the physician believes the profile is an accurate reflection of his practice. Several checklists for critical review of physician performance profiles have been published [14–16]. We highlight five types of methodologic issues that are central to judging the quality of physician profiles (**Table 2**).

Representativeness

To be valid, data on physician performance must include a *representative sample* of the physician's patients. For example, the Web sites of several Internet-based companies we reviewed report patient satisfaction scores for individual physicians based on information solicited from patients through the Web site itself. Patients who answer satisfaction questions on a Web site clearly are not representative of all patients in a physician's practice. On the other hand, profiles that report all patients with a given condition seen by a physician are by definition representative of the patients in the physician's practice. Patient satisfaction surveys as well as practice-based measures (eg, the percentage of eligible women who received a Pap smear) should include a representative sample of patients encountered by the physician or practice being profiled.

Accuracy

Profiles can be based on many different data sources, such as claims data (ie, the information given to the insurance company or MCO as a basis for reimbursement) or patient or physician surveys. Because different data sources vary in their ability to reflect actual practice, the accuracy of the information used in profiles should be verified through direct measurement, such as review of administrative or medical records. For example, the Internet company DoctorQuality asks physicians to report their compliance with HEDIS measures by responding with a "yes" or "no" answer. The Web site does not report the accuracy of the responses, which could be determined by examining the medical records of a sample of respondents.

(continued on page 30)

Table 1. Examples of Organizations that Release Physician Profiles to Public Audiences

Profiling Organization	Intended Audience	Measures
Managed Care Organizations*		
Health Partners, Minnesota (www.healthpartners.com)	Patients eligible or enrolled in the MCO	Patient satisfaction with several measures of care Provider performance on 11 measures
PacifiCare, California (www.pacificare.com)	Patients eligible or enrolled in the MCO	25 measures of process of care and patient satisfaction
State Governments		
Pennsylvania Health Care Cost Containment Council (www.phc4.org/reports/cabg9495/default.htm)	Health care consumers and physicians	CABG procedure volumes CABG mortality rate
New York State Department of Health (www.health.state.ny.us/homens.html)	Health care consumers and physicians	CABG procedure volumes CABG mortality rate
Federal Government Agency		
Center for Medicare and Medicaid Services (CMS)† (www.chms.hhs.gov)	Physicians	Physician-specific performance data for Medicare patients with diabetes as part of the 6th Scope of Work
Health Care Coalitions		
Pacific Business Group on Health (www.pbgh.org)	Health care purchasers and consumers	Patient satisfaction
Greater Detroit Area Health Council (www.gdahc.org/Csig/cardiac.htm)	Health care purchasers and consumers	Volume of cardiac diagnostic and surgical procedures
Central Florida Health Care Coalition	128 member businesses Health care providers, to inform quality improvement	Commercially produced performance measures based on clinical and administrative data
Nonprofit Patient Advocacy Groups		
American Diabetes Association (ADA) (www.diabetes.org/main/professional/recognition/default.jsp)	Diabetic patients	Evidenced-based measures consistent with those of the Diabetes Quality Improvement Project
Internet Companies		
Doctor Quality (www.doctorquality.com)	Health care consumers and purchasers	Patient satisfaction; self-reported physician disciplinary actions; HEDIS measures
Health Grades (www.healthgrades.com)	Health care consumers and purchasers	Patient satisfaction
Health Care Choices (www.healthcarechoices.org)	Health care consumers and purchasers	Surgical procedure volumes for breast cancer surgery for physicians in New York State
Best Doctors (www.bestdoctors.com) (Other services: FindBestDoc, AcuMatch)	Health care consumers	No measures, just names of recommended physician(s)

NOTE: The information in this table was gathered by nonsystematic survey and is not intended to represent all entities currently producing physician performance profiles.

CABG = coronary artery bypass graft; HEDIS = Health Plan Employer Data and Information Set; MCO = managed care organization; PPO = preferred provider organization; PRO = peer review organization.

*Describes published profiles only.

†Formerly the Health Care Financing Administration.

(*Table 1 continued on next page*)

	Data Collection Method	Comments
	Patient surveys and patient medical records	Performance profiles published for medical groups, not individual physicians
	Records submitted by medical group and patient satisfaction surveys	Performance profiles published for medical groups, not individual physicians
	Patient medical records	Developed highly detailed risk adjustment system for reporting mortality
	Patient medical records	Developed highly detailed risk adjustment system for reporting mortality
	Administrative data	Political and economic pressure to release individual performance information to health care providers and patients will remain high
	Patient satisfaction surveys	Profiles performance of medical groups; plans to profile medical groups using HEDIS process of care measures for asthma and diabetes
	Volume for each physician was voluntarily reported by hospitals	Active in performance measurement; promotes efforts of local PROs to measure and improve diabetes care by individual physicians
	Data uniformly collected through common software by member physician groups and hospitals	Viewed as a front-runner in linking payment and work incentives directly to individual physician performance
	Patient medical records and surveys Physicians submit data to ADA and pay application and program fees	Voluntary program for physicians; cosponsored by NCQA; physicians must have minimum of 35 diabetes patients in a 12-month period.
	Patients complete Web-based satisfaction survey; physician self-report of disciplinary actions and qualitative responses for HEDIS measures	Established an agreement with a Wisconsin PPO to report measures for network physicians
	Patients complete Web-based satisfaction survey	Provides information on board certification, health plan hospital affiliation, sanction activity, and accepting new patients to employers and MCOs
	New York State Department of Health	Also provides links to state physician profile Web sites and New York and Pennsylvania Web sites on individual surgeon performance for CABG
	Surveys of approximately 30,000 physicians in 40 different specialties	A fee is charged for FindBestDoc and AcuMatch; AcuMatch involves a search for specialized physicians for treatment of serious or complex medical problems and includes services of a case manager

(continued from page 27)

Physician and hospital billing (ie, claims) and encounter records are the most commonly used information sources for physician profiling. In general, these administrative data are accurate (condition present if recorded), but information on important chronic diseases or comorbidities may not be captured. For example, claims data may record hypertension as the reason for a patient's visit but not record that the patient also has diabetes and depression and uses tobacco. In general, profiling efforts that capture a large number of claims or encounters for a given patient will be more likely to include all relevant clinical conditions for that patient.

Medical records are highly specific (condition present if recorded) and sensitive (condition not present if not recorded) for some data such as immunizations and Pap smears but are insensitive for other information such as counseling [17] and psychosocial conditions [18]. Medical record review can be useful for validation studies of administrative and claims data, but this method is relatively labor-intensive and therefore too costly to be used as an information source for ongoing profiling efforts.

Reliability

Reliability is the extent to which a measure gives the same answer on repeated trials. It is related to the number of observations available for an individual physician and to the amount of practice variation between physicians for a particular measure. Reliability increases as the number of observations for each physician and the practice variation between physicians increase. To illustrate this relationship, consider the example of measuring the proportion of patients with diabetes in a physician's practice who receive annual eye examinations. If a given physician treated only one patient with diabetes, information about that patient's care is representative (100% sample) and may be accurate (good source data), but it is not reliable. That is, we know that the physician's care of that single patient will not allow us to predict with any certainty how he will care for the next patient with diabetes. Next, consider two physicians in the same practice, each of whom cares for 100 patients with diabetes. If only 20% of physician A's patients receive annual eye examinations while 80% of physician B's patients receive examinations, a smaller number of patients from each physician's practice must be sampled to accurately detect the difference between the two physicians than if their actual referral rates were more similar, for example 40% and 45%.

Reliability is expressed as a value between 0 and 1. A reliability of 0.8, which indicates that 80% of the variation in a physician's profile is due to practice differences and 20% is due to chance variation, has been considered the minimum level of reliability necessary for making decisions about an individual physician based on a profile [5]. For chronic conditions, it is extremely difficult to achieve reliability at the level of the individual physician. For example, Hofer et al [5] studied the performance of several large primary care practices in managing patients with type 2 diabetes and found that the reliability of a measure of glycemic control based on glycosylated hemoglobin levels was only 0.38 [5]. The average number of diabetic patients per physician studied was 21, but a panel size of at least 100 would have been required to achieve adequate reliability. To place this finding in context, consider that the American Diabetes Association (ADA) has initiated a program designed to recognize physicians who manage patients with diabetes well. The ADA Provider Recognition Program, however, uses similar measures to those used in the study by Hofer and colleagues [5] but requires only 35 sample patients. Thus the ADA Provider Recognition Program is unlikely to provide reliable estimates of an individual physician's performance in managing patients with diabetes.

The sampling requirements needed to produce reliable profiles for individual physicians are formidable. However, some have argued that for certain conditions (eg, mortality following CABG), the ability to capture all relevant outcome data (eg, deaths) and the existence of standardized interventions (eg, CABG) will allow reliable profiles of individual physicians to be created with relatively small samples [19]. The feasibility of this approach, however, has not been demonstrated.

Attribution

Care and patient characteristics must be correctly attributed to the physician being profiled. This is fairly straightforward when measuring a surgeon's clinical outcomes following CABG but is more difficult in other circumstances, such as measuring preventive care services for patients who receive primary care from several different physicians. In such cases, it is more appropriate to profile the group practice rather than the individual physician.

Case-mix Adjustment

Case-mix adjustment is a method used to control for patient characteristics that affect outcomes of care (eg, age, sex, comorbid conditions, and severity of disease) and is crucial to allow fair comparisons of the practice profiles of individual or groups of physicians [20–23]. Case-mix adjustment is particularly important in profiling resource use (eg, pharmacy costs and utilization,

inpatient costs and utilization) and is less important when measuring processes of care (eg, mammography rates among women 50 to 74 years of age) [24]. Note that in the opening scenario, Dr. Schroeder does not need to consider case-mix adjustment for the two process measures as long as the patients included in the report are her patients.

Conceptually, case-mix adjustment occurs in three steps. First, a classification system is developed that allows prediction of the resource use of individual or groups of patients based on their demographic, clinical, or health services utilization characteristics. These systems are usually complex and are developed and validated through research on large administrative data sets. Somewhat confusingly, the classification systems are often called case-mix adjustment systems (or methods). Although several systems have been developed, the two most widely used and well-validated are the Adjusted Clinical Groups (ACG) and Diagnostic Cost Groups (DCG) systems. After classifying individual or groups of patients, an index number or weight is assigned that reflects the patient's or group's *risk* of resource consumption. Finally, the measured resource in physician profiles are weighted (*adjusted*) by the index reflecting the case-mix characteristics of the patients included in the profile.

Physicians should know which classification system and weighting method, if any, are applied to profiles reflecting their practice. The ACG system uses claim and encounter experience over 6 to 12 months to assign each patient to a single ACG. The ACG system is used to predict overall health care costs or costs within classes of services (eg, inpatient care and ambulatory visits, laboratory tests) for groups of patients [24]. It is not designed to predict costs for individual patients. Although the ability of case-mix adjustment systems to predict resource use diminishes substantially for populations of fewer than 2000 patients, ACGs were developed to minimize bias and can arguably be used to predict resource use for as few as 250 patients [24]. ACGs have been demonstrated to be superior to other existing case-mix adjustment systems in describing the actual cost profiles of physicians in ambulatory practice [25].

The DCG system is a family of profiling methods that uses the demographics and diagnoses recorded on claim and encounter data to predict the relative resource use of individual patients. Unlike ACGs, which attempt only to project total costs for populations, the DCGs also attempt to predict total costs for individual patients. DCG systems have been developed for Medicare, Medicaid, and privately insured patients [24].

Patients whose resource use is significantly higher than predicted are termed *outliers*, and these patients present a challenge to any case-mix adjustment system. That is, case-mix–adjusted physician profiles generally improve dramatically by eliminating a few (eg, one to three) patients with poor profiles. This provides incentives for physicians to "game" the system by choosing not to treat difficult or challenging patients [5]. The complex algorithms of the ACG and DCG systems were developed specifically to minimize the potential for this type of gaming.

No case-mix adjustment method is perfect, and ongoing research is being conducted to create new or refined classification systems [26]. Where possible, physician profiles should focus on conditions that are relatively insensitive to comorbidity and severity of illness. For example, measuring rates of referral for annual eye examinations among diabetics, a service that arguably every diabetic patient should receive, is preferable to measuring pharmaceutical costs for a given physician's patients with diabetes, which may vary greatly depending on comorbidities and severity of illness.

Table 2. Checklist for Methodologic Review of Physician Profiles

Representativeness. Patients included in the profile were fairly sampled from the target group in the physician's practice.
Reliability. Adequate number of patients were included.
Accuracy. Data source has been demonstrated to have acceptable accuracy.
Attribution. The activity being profiled for a given group of patients can be fairly attributed to the physician or physicians being profiled.
Case-mix adjustment. A validated case-mix adjustment system was used to account for differences in comorbidities or severity of illness that can result in differences in resource use.

What is the Prognosis for Methodologic Problems?

The scope and severity of methodologic problems in physician profiling are just beginning to be understood and addressed. At the same time, the practice of profiling is growing rapidly to meet the need for information on the quality and amount of care provided by physicians. Physician profiling is here to stay [15,27,28]. Methodologically sound profiling efforts, however, will find their greatest application in describing the practices of relatively large groups of physicians due to the formidable sampling and reliability problems in creating profiles for small groups or individual physicians. Their

utility in describing the practices of individual physicians, especially for high-stakes applications (eg, public release or pay incentive systems targeted at individual physicians) remains limited for the foreseeable future [1].

How has Profiling Impacted Physician and Patient Behavior?

Because physician profiles serve many purposes, the intended audiences may include physicians, patients, health systems, and employers. The effects of profiling on health systems and health plans are beyond the scope of this article but are the subject of ongoing research [29–31]. This discussion focuses on the impact of physician profiling on physician and patient behavior.

Several recent meta-analyses systematically reviewed randomized trials to quantitatively estimate the effects of profiling and related activities on physician behavior and clinical outcomes [2,32–34]. In these studies, profiling was defined as any systematic collection of representative information on physician behavior (also called *audit*) and provision of that information to the physicians being described (termed *feedback*). Target behaviors included diagnostic test ordering, prescribing practices, preventive care, and the general management of a clinical problem (eg, hypertension). In general, these meta-analyses found that audit with feedback has a small but significant effect on physician behavior, especially when combined with educational materials such as clinical guidelines or literature summaries that informed physicians of best practices in managing particular disorders (eg, asthma, diabetes, depression). Further improvement in physician behavior was seen when reminders for preventive care services were added. More recent studies are consistent with the findings from these meta-analyses [27,35,36]. Additional elements of successful strategies include targeting specific objectives for improvement, rather than providing profiles without context [37], and involving the target physicians in the creation, dissemination, and application of profiles [38].

Several agencies and organizations make physician profiles available directly to consumers (Table 1), ostensibly as an aid in making informed decisions on choosing or continuing care with particular physicians. There is considerable evidence, however, that consumers are unlikely to rely on published performance profiles to choose a provider [29,39] and instead are more likely to rely on the anecdotes and recommendations of friends and family [10]. Barriers to the use of profiles by consumers include comprehensibility, perceived importance, and level of concreteness [10,40]. Some researchers have concluded that consumers require education to interpret profile reports, especially in instances where the reports display measurement results that are not uniformly favorable [40].

What Current Profiling Trends are Likely to Continue?

To quote Epstein [28], "The genie is out of the bottle." Clearly, profiling is going to continue; less clear are the direction and focus of future physician profiling efforts. We anticipate that at least two trends in profiling will continue to increase: 1) physician involvement in profiling, and 2) use of profiles as a basis for performance-based payment to physicians.

Physician organizations have been slow to become involved in physician profiling, often taking the contrarian position that profiles are not valid due to methodologic problems [4]. This approach has limited the involvement of physician organizations in the creation and application of meaningful information-based feedback systems to improve the quality and efficiency of medical practice [15]. Opportunities for physicians to measure and improve the quality and cost of care at all levels will continue to grow—in group practices, within insurance organizations or HMOs, in integrated health systems (eg, hospitals, physician practices, and ancillary services), and in regional broad-based efforts to improve health care and limit costs (eg, Pacific Business Group on Health, Detroit Area Health Council). Most common will be opportunities for physicians to constructively engage in measuring and improving their own practices. Active physician involvement in the creation and translation of profile information into quality improvement activities is likely to lead to more accurate, relevant, and productive profiles [4,12,28].

It is estimated that incentives based on quality of care are offered to 19% of physicians [41]. We anticipate that interest will remain high in using physician profiles as a means for making performance-based payments for quality care. Use of performance profiles for this type of high-stakes application will likely grow slowly, however, due to the formidable methodologic challenges in creating fair and objective reports for individual physicians. For example, some organizations such as JCAHO and NCQA plan to avoid them (Kathy Berry, JCAHO, personal communication, 26 July 2001) or approach them very cautiously (Brian Schiller, NCQA, personal communication, 3 July 2001). Other groups such as employer groups and regional consortia of businesses, health plans, and providers will likely continue efforts to develop and use profiles as a basis for rewards and penalties to individual physicians. For example, the Central

Florida Health Care Coalition, a group of 128 businesses, health plans, and providers, is moving forward to develop profile-based reward systems for individual physicians [42].

Summary

Physician profiling by a variety of organizations is on the rise. Moreover, as Eisenberg noted in a recent editorial [43], as providers and consumers of health care become more comfortable with performance measures and their reliability, "we will probably see them being used not just by purchasers who want to make value-based decisions but also by providers who want to improve the care in their practice; demonstrate that they give excellent care; and provide the public with the accountability and informed, value-based choice that it wants." The concepts of physician profiling—the purpose, content, and methodologic issues in producing meaningful profiles—are not complex. Furthermore, the skills necessary to apply, develop, or improve profiles are well within the reach of practicing physicians and can be applied in systematic steps (**Appendix**). Physicians should participate productively in the profiling process and take advantage of the opportunities presented by profiling for improving patient care and practice management.

Address correspondence to: Vinita Bahl, DMD, MPP, Director of Clinical Information and Decision Support Services, University of Michigan Hospitals and Health Centers, 300 North Ingalls Building, Room 7A10, Ann Arbor, MI 48109-0485 (e-mail: Vbahl@umich.edu).

References

1. Bindman AB. Can physician profiles be trusted? JAMA 1999;281:2142–3.
2. Balas EA, Boren SA, Brown GD, et al. Effect of physician profiling on utilization. Meta-analysis of randomized clinical trials. J Gen Intern Med 1996;11:584–90.
3. Ullman R, Scherpbier H, Diamond C. Physician profiling in managed care. Sem Med Pract 1999:14–20.
4. Newcomer LN. Physician, measure thyself. Health Aff 1998;17:32–5.
5. Hofer TP, Hayward RA, Greenfield S, et al. The unreliability of individual physician "report cards" for assessing the costs and quality of care of a chronic disease. JAMA 1999;281:2098–105.
6. Emmons DW, Wozniak GD. Profiles and feedback: who measures physician performance? In: Gonzalez ML, editor. Socioeconomic characteristics of medical practice. Chicago (IL): American Medical Association; 1994.
7. Goldfield N. Harvesting data for physician profiling. In: Piland NF, Lynam KB, editors. Physician profiling: a source book for health care administrators. San Francisco: Jossey-Bass, Inc.; 1999:3:31–46.
8. Kerr EA, Mittman BS, Hays RD, et al. Managed care and capitation in California: how do physicians at financial risk control their own utilization? Ann Intern Med 1995;123:500–4.
9. Lawthers AG, Palmer RH. In search of a few good performance measures: CONQUEST and the typology of clinical performance measures. In: Seltzer J, Nash DB, editors. Models for measuring quality in managed care: analysis and impact. New York: Faulkner & Gray Healthcare Information Center; 1997:121–50.
10. Hibbard JH, Jewett JJ. Will quality report cards help consumers? Health Aff 1997;16:218–28.
11. Jackson C. California HMO: doctor bonuses based on patient satisfaction. American Medical News 30 July 2001.
12. Schneider EC, Epstein AM. Use of public performance reports: a survey of patients undergoing cardiac surgery. JAMA 1998;279:1638–42.
13. Green J, Wintfeld N. Reports on cardiac surgeons. Assessing New York State's approach. N Engl J Med 1995;332:1229–33.
14. Bell KM. Physician profiling: 12 critical points. J Ambulatory Care Manage 1996;19:81–5.
15. Tran ZV, Burman DL. Introduction: The story of profiling. In: Piland NF, Lynam KB, editors. Physician profiling: a source book for health care administrators. San Francisco: Jossey-Bass, Inc.; 1999:xix–xxxv.
16. American Association of Family Practice. Physician profiles guiding principles. AAFP reprint No. 205. March 1999. (www.aafp.org).
17. Stange KC, Zyzanski SJ, Smith TF, et al. How valid are medical records and patient questionnaires for physician profilings and health services research? A comparison with direct observation of patient visits. Med Care 1998; 36:851–67.
18. Williams BC, Philbrick JT, Becker DM, et al. A patient-based system for describing ambulatory medicine practices using diagnosis clusters. J Gen Intern Med 1991;6: 57–63.
19. Scanlon DP, Chernew M, Doty HE, Smith DG. Options for assessing PPO quality: accreditation and profiling as accountability strategies. Med Care Res Rev 2001b;58 Supplement 1:70–100.
20. Goldfield N. Understanding your managed care practice: the critical role of case mix systems. Manag Care Q 1994;2:12–20.
21. Salem-Schatz S, Moore G, Rucker M, Pearson SD. The case for case-mix adjustment in practice profiling. When good apples look bad. JAMA 1994;272:871–4.
22. Kolb DS, Clay SB. Ambulatory care groupings: when, how, and the impact on managed care. J Ambulatory Care Manage 1994;17:29–38.
23. American Medical Association. Physician Consortium for Performance Improvement. Available at http://www.ama-assn.org/ama/pub/category/2946.html. Accessed 5 May 2002.
24. Ruben MS, Braun P, Caper P. Case mix adjusters. In: Piland NF, Lynam KB, editors. Physician profiling: a

source book for health care administrators. San Francisco: Jossey-Bass, Inc.; 1999:17–30.
25. Tucker AM, Weiner JP, Honigfeld S, Parton RA. Profiling primary care physician resource use: examining the application of case mix adjustment. J Ambulatory Care Manage 1996;19:60–80.
26. Greenfield S, Kaplan SH, Kahn R, et al. Profiling care provided by different groups of physicians: effect of patient case-mix (bias) and physician-level clustering on quality assessment results. Ann Intern Med 2002;136:111–21.
27. Weiss KB, Wagner R. Performance measurement through audit, feedback, and profiling as tools for improving clinical care. Chest 2000;118(2 Suppl):53S–58S.
28. Epstein AM. Rolling down the runway: the challenges ahead for quality report cards. JAMA 1998;279:1691–6.
29. Bentley JM, Nash DB. How Pennsylvania hospitals have responded to publicly released reports on coronary artery bypass graft surgery. Jt Comm J Qual Improv 1998;24:40–9.
30. Romano PS, Rainwater JA, Antonius D. Grading the graders: how hospitals in California and New York perceive and interpret their report cards. Med Care 1999; 37:295–305.
31. Scanlon DP, Darby C, Rolph E, Doty HE. The role of performance measures for improving quality in managed care organizations. Health Serv Res 2001;36:619–41.
32. Davis DA, Thomson MA, Oxman AD, Haynes RB. Changing physician performance. A systematic review of the effect of continuing medical education strategies. JAMA 1995;274:700–5.
33. O'Brien MAT, Oxman AD, Davis DA, et al. Audit and feedback: effects on professional practice and health care outcomes. In: The Cochrane Library; 1, 2002. Oxford: Software Update.
34. O'Brien MAT, Oxman AD, Davis DA, Haynes RB, et al. Audit and feedback versus alternative strategies: effects on professional practice and health care outcomes. In: The Cochrane Library; 1, 2002. Oxford: Software Update.
35. Preston JA, Scinto JD, Grady JN, et al. The effect of a multifaceted physician office-based intervention on older women's mammography use. J Am Geriatr Soc 2000;48:1–7.
36. Lagerlov P, Loeb M, Andrew M, Hjortdahl P. Improving doctors' prescribing behaviour through reflection on guidelines and prescription feedback: a randomised controlled study. Qual Health Care 2000;9:159–65.
37. Boren SA, Balas EA, Su KC. Seeing the forest through the trees. In: Piland NF, Lynam KB, editors. Physician profiling: a source book for health care administrators. San Francisco: Jossey-Bass Inc.; 1999:47–64.
38. Smith NS. Laying the foundation for a profiling system. In: Piland NF, Lynam K, editors. Physician profiling: a source book for health care administrators. San Francisco: Jossey-Bass, Inc.;1999:3–16.
39. Schneider EC, Epstein AM. Influence of cardiac-surgery performance reports on referral practices and access to care. A survey of cardiovascular specialists. N Engl J Med 1996;335:251–6.
40. Hibbard JH, Slovic P, Jewett JJ. Informing consumer decisions in health care: implications from decision-making research. Milbank Q 1997;75:395–414.
41. Few docs subject to incentives to limit care, says HSC. Medicine & Health 11 February 2002.
42. A conversation with Becky Cherney. Managed Care Magazine December 2000. Available at http://www.managedcaremag.com/archives/0012/0012.qna_cherney.html. Accessed 5 May 2002.
43. Eisenberg J. Measuring quality: are we ready to compare the quality of care among physician groups? Ann Intern Med 2002;136:153–4.

Appendix. Recommended Physician Approach to Profiling

Become knowledgeable about physician profiles in general, particularly:

- General trends in profiling (eg, through articles, journals, newsletters, and Internet-based news sources)
- Performance measures that are widely used (eg, Health Plan Employer Data and Information Set [HEDIS] clinical performance measures and Consumer Assessment of Health Plans [CAHPS] patient satisfaction measures)
- High-quality case-mix classification systems likely to be used in profiling (eg, Adjusted Clinical Groups [ACG], Diagnostic Cost Groups [DCG])
- Goals and activities of organizations active in developing methods for measuring performance (eg, National Committee for Quality Assurance [NCQA], Foundation for Accountability [FACCT], Joint Commission on Accreditation of Healthcare Organizations [JCAHO], American Medical Association)

Review your attitudes toward physician profiling.

- Negative attitudes and nonparticipation are unproductive.
- Informed, constructive attitudes toward profiling will help your practice.

Determine how profiling is currently used in your practice.

- Sponsoring agency. Are profiles being created by payers, employers, or providers?
- Purpose. Will the profile be used for a high-stakes (eg, payment) or a low-stakes (eg, feedback for reflection) application?
- Content. Does the profile focus on resource consumption, clinical performance, patient perceptions, or all of these?
- History. How did profiling begin in your practice? Who initiated it? What were the conditions that created the perceived need for profiles?

Critically review existing practice profiles for content and quality.

- Review existing profiles for methodologic soundness (ie, representativeness, accuracy, reliability, attribution, case-mix adjustment)
- Determine the population of patients included in the profile. Is the population (denominator) well defined? Does the population represent an adequate sample of your total practice?
- Comparison group. To what group are the profile results compared as the norm or target? Does the comparison group represent practices and patients roughly similar to yours?

Choose a few, selected profiling measures to address. It is not feasible to address all actual or potential areas for practice improvement at once. Choose performance measures that:

- Are applicable in multiple contexts. For example, HEDIS measures are likely to be useful for both internal quality improvement and external review purposes.
- Represent prevalent and clinically important conditions among patients and across providers.
- Reflect physician behavior. Focus on clinical measures (eg, mammography rates) rather than nonclinical (eg, wait times, parking) measures.
- Link clinical decisions to patient outcomes. For example, the use of angiotensin-converting enzyme inhibitors in hypertensive diabetics has been shown to improve outcomes, whereas nutrition counseling for all overweight patients has not.

Gather or obtain additional information to follow up the profile results.

- Profiles often raise the possibility of quality or cost problems but may not fully demonstrate their presence or delineate the causes of the problems. Managed care organizations and hospitals are often interested in providing such additional data to interested practices.

Apply profiling information to improve your practice.

- Organize multifaceted quality improvement activities around the end points reported in existing profiles. Physician review of profile information alone is unlikely to result in meaningful change. Combining performance profiling with other interventions such as education, reminders, and modification of the practice organization (eg, facilitating access to influenza vaccination or mammography outside physician office visits) are more likely to be effective.
- Incorporate the profile results into continuous quality improvement activities. The processes of continuous quality improvement—Plan (set specific objectives around areas in demonstrated need of improvement), Do (implement quality improvement interventions), Check (measure the effects of the interventions), Act (modify objectives, intervention, or data gathering methods for improvement)—describe a systematic approach to quality improvement that has been widely advocated in health care. The data provided by performance profiles can serve as one component of a broader set of quality improvement activities [1,2].

(continued on next page)

Appendix. (*continued*)

If needed, work with existing profiling agencies to improve the usefulness of physician profiles.

- If existing profiles cannot be used as one component of systems to improve the quality or manage the costs of your practice, or if the quality of the profiles prohibits meaningful interpretation, they cannot be used productively by you or the practice. Consider working directly with the profiling agency (managed care organization, hospital, accrediting organization) to design or modify the profiling system to provide more useful information.

Identify opportunities to gather new performance information to improve the quality of your practice.

- Determine the prevalence of diseases or diagnoses in your practice to identify types of problems that warrant improvement efforts.
- Identify likely quality or cost problems. Are they likely to be related to the practice patterns of individual physicians? Further information gathering will often be necessary to determine whether a problem exists.
- Identify existing information sources that lend themselves to performance profiling. Existing administrative data may be available to document key processes of care. The most useful markers of quality are often processes of care that require little case-mix adjustment, such as preventive health care practices (eg, immunization or mammography rates) and care that is routine, diagnosis-specific, and evidence-based (eg, annual eye and foot examinations for patients with diabetes).

References

1. American Association of Family Practice. Physician profiles guiding principles. AAFP reprint No. 205. March 1999.
2. Scanlon DP, Darby C, Rolph E, Doty H. Use of performance information for quality improvement: the role of performance measures for improving quality in managed care organizations. Health Serv Res 2001;36: 619–41.