

PHYSICIAN PERFORMANCE MEASUREMENT: A KEY TO HIGHER QUALITY AND LOWER COST GROWTH OR A LOST OPPORTUNITY?

By Debra A. Draper

Although the United States spends more than \$2 trillion annually on health care, patient outcomes lag other developed countries that spend far less per capita. Physicians wield significant influence—directly and indirectly—over the quality and cost of health care, and efforts to measure and improve physician performance have gained momentum. Much of the impetus has come from purchasers seeking to engage consumers to be more active participants in their health and health care decisions. In response, health plans have developed physician performance measurement programs to provide information to consumers. However, methodological limitations, including the use of claims data, small sample sizes, and non-standardized measures and assessments, have fueled skepticism about plan programs. While measuring performance is an important step, health plans often fail to take the next step—supporting and rewarding physician performance improvement to encourage and reinforce desired behaviors. Arguably, physician performance measurement has such profound implications for all Americans' health and health care that it should be a public good, transcending competitive dynamics. Standardizing measures, combining payers' data, providing effective support for improvement, and creating robust rewards for good results offer some ways to improve the current state of physician performance measurement.

Physicians Key to Higher Quality and Lower Cost Growth

U.S. health care costs continue to spiral upward. In 2007, the United States spent \$2.2 trillion on health care, or 16 percent of the nation's gross domestic product (GDP), and spending grew more than 6 percent from the previous year.¹ Yet, despite the highest per-capita health expenditures in the world, U.S. patient outcomes are comparatively worse than those of many other developed countries with much lower spending.² The disconnect between money spent on health care and the often less-than-stellar results has sparked national awareness of the critical importance of measuring and improving health care quality and slowing spending growth through increased efficiency. As a result, nascent efforts are underway to measure and improve physician performance on both quality and cost dimensions.

Physicians are the linchpin of care delivery, and, directly and indirectly, they have significant influence on health care quality

and costs.³ Measuring physician performance to identify weaknesses that warrant change and working to make those changes, therefore, creates tremendous opportunity to improve health care quality and efficiency. Although physician performance measurement and improvement offers a potentially powerful tool, it may prove a lost opportunity for improving the nation's health care system if methodological and other shortcomings of existing efforts are not appropriately addressed.

Efforts to Measure and Reward Physician Performance

To date, most performance measurement programs have been developed by health plans seeking to differentiate physicians on the basis of quality and costs. Much of the impetus has come from purchasers, notably large national employers, hoping to address

quality and cost concerns by engaging consumers to be more active participants in decisions about their health and health care.⁴ As the responsibility for health care decision making and costs increasingly shifts to consumers, there is a recognized need to provide more and better information about health care providers, including the quality and cost-effectiveness of the services these providers deliver.⁵ Plans have embraced these measurement efforts as a way of creating value for their employer clients and to help distinguish themselves in a competitive marketplace.

Plan efforts often are manifest and marketed in the form of physician ranking programs or some type of narrow, tiered or high-performance provider network.⁶ These programs operate under a variety of names such as the Aexcel Specialist Network (Aetna), Blue Precision (Blue Cross Blue Shield), Care Network (CIGNA), Preferred

Network (Humana), and Premium Designation Program (UnitedHealthcare). The underlying premise of these initiatives is to provide a systematic and objective method of measuring physician performance based on quality and cost metrics that can be assessed using plans' claims or other administrative data and making the results publicly available to enrollees. Most often, the results are used only to inform consumers; in some cases, consumers have incentives, such as reduced copayments, to use the higher-performing physicians. Plans rarely pay bonuses to physicians they deem high performing. In these programs, quality and efficiency improvements are achieved to the extent that patient volume shifts to higher-performing physicians as a result of changes in physician referrals and consumer choices and lower-performing physicians improving the care they provide.

Although plans' physician performance measurement programs are broadly similar, they vary in the methodologies employed. Methodologies often differ on dimensions such as the specific measures used, sample-size requirements, and the comparative emphasis placed on quality vs. cost measures. Consequently, gauging the comparability of individual plan results is difficult because the decision algorithm each plan uses to conduct the assessments is proprietary with little, if any, transparency. This variability can result in physicians deemed high performing by one plan but not another, as was the case, for example, for a large integrated delivery system in Seattle, Virginia Mason Medical Center, as plans rolled out their respective programs in that market.⁷

Limited physician input and lack of transparency, which the American Medical Association describes as "black-box methodologies," has resulted in considerable physician skepticism and outright dismissal by some. It also has resulted in legal action. In 2006, for example, the Washington State Medical Association filed suit against Regence Blue Shield, alleging Regence used flawed methods and outdated information to exclude physicians from the plan's high-performance network. The pushback resulted in Regence discontinuing the program, at least until it could be revamped.⁸ In 2007, New York Attorney General Andrew

Cuomo launched an investigation into the physician ranking programs of health plans operating in New York, raising concerns that plans' profit motives affected the accuracy of the rankings and encouraged consumers to choose physicians solely on the basis of cost.⁹ As a result of the investigation, health plans agreed to make a number of changes, including basing their assessments not solely on costs, using national quality and efficiency measures, and using measures that help facilitate consumers' comparisons of physicians. The agreement also required plans to score 100 percent compliance on external reviews of their ranking programs.¹⁰

As these reactions to health plans' programs suggest, performance measurement can evoke anxiety generally and be especially threatening to physicians who are unlikely to show good performance. The stakes are high for physicians deemed poor performing, because their professional reputations are at risk and, potentially, their financial interests. Ensuring that a valid methodological approach is used to measure performance is therefore crucial because egregious, albeit unintended, consequences could include incorrectly labeling a physician as a poor performer or having a consumer choose a physician based on an inaccurate assessment whose care resulted in an adverse outcome. As recent history suggests, methodologies to assess physician performance are subject to intense scrutiny, and weaknesses in objectivity, credibility and transparency can undermine, if not derail, the intended objectives of improving health care quality and efficiency. Engaging physicians as active participants in plans' performance improvement efforts has proved difficult because of these weaknesses. And when there are problems with one plan's performance measurement effort, physicians often construe problems more broadly to all plan efforts.

Methodological Shortcomings Tarnish Credibility

Although strong methodological approaches to physician performance measurement are vital to its success, shortcomings have tarnished—at least initially—the credibility

of many health plans' efforts. Much of the controversy has focused on data credibility, sample sizes and methods used to analyze the data.

Lack of Data Credibility. Plans typically use their own claims and other administrative data to measure physician performance. However, these data can be considerably less reliable and accurate than data extracted through medical record review, which is more expensive to collect. Claims and administrative data have inherent weaknesses in documenting all services provided to a patient by a physician and in capturing legitimate reasons why certain services were or were not provided—information that is critical for an accurate assessment of physician performance. Typically, health plans do not collaborate with other entities, for example, a large physician group with robust electronic medical record data, to compare and validate claims data and factor in any needed data adjustments.

Inadequate Sample Size. Although plans typically require a minimum sample size to assess a physician's performance, these thresholds tend to be set relatively low (e.g., fewer than a dozen patients) in part because of limitations associated with an individual plan's use of only their own claims data to conduct the assessment. But because any single plan's patients may represent only a small fraction of a physician's entire patient panel, there is a greater likelihood that the assessment may yield incomplete, if not, erroneous results. For example, if a plan's patients are disproportionately sicker with higher costs of care than the physician's overall patient panel, the plan's assessment might tag the physician as a poor performer, when the opposite may be true.

Non-standardized Measures. There is no standardized set of measures used by plans to assess physician performance, and even if the measures are the same or similar, plans may define and operationalize them differently. The same is true of methods to adjust for differences in risk among patients—whether and how these methods are applied to the measures. Adjustment for patient differences is important because many physicians believe their patients are more challenging

than average. Plans typically use evidence-based medical guidelines and consensus-based quality standards to assess physician quality. Efficiency is generally measured using episodes of care and attributing all related costs, including those of other providers, to the physician deemed primarily responsible for the patient's care, regardless of whether that physician has control over the other providers rendering the care. Physician organizations have been critical of the tools—called groupers—used to sort claims into episodes of care in part because this methodology is still evolving, but also because of the way in which physicians are assigned and held accountable for all of the costs of a patient's care.

Non-standardized Assessment. There is little, if any, consensus among plans about how physicians' performance should be assessed, including, for example, the relative emphasis of quality vs. cost measures. This is at least partially rooted in plans' desire to use their physician performance measurement efforts as a way of gaining a competitive advantage in the marketplace—to have something different and seemingly better than their competitors to offer employer clients. The difficulty with such a proprietary approach, however, is that it creates distrust because of the limited transparency of the process and how the results were derived. It also diminishes the overall credibility and effectiveness of the assessments because there is no comparability between plans, leading to physicians deemed high performing by one plan but not another and creating confusion for those that may rely on the information.

Measurement Necessary but Not Sufficient

Measuring performance is an important and necessary initial step to improving the quality and efficiency of care provided, but measurement alone is likely insufficient to prompt physicians to improve performance. Support to improve performance and rewards to encourage and reinforce desired behaviors also are needed. However, these elements are largely absent from many health plan efforts, and when they do exist, they are inadequate to bring about meaningful or sustained perfor-

mance improvement.¹¹

Support for Improving Performance.

It is of little use to measure performance and not also support physicians willing to improve. Health plan assessment efforts generate considerable data, but many fail to provide physicians the information in a clear and actionable manner. A case study of Virginia Mason Medical Center is illustrative in understanding the value of providing meaningful data to foster performance improvement.¹² In this particular case, Aetna provided detailed claims data by individual physicians, practice sites, patients and cost centers, such as pharmaceuticals and emergency services, which then allowed the system to conduct further analyses to identify cost-reduction opportunities. Through this process, for example, Virginia Mason identified that its costs per migraine episode were high, in part because patients went to the emergency department for severe headaches when they lacked "rescue" medication. The analysis provided the system with important information to help guide physicians in changing their care of patients with migraines.

Benchmarking an individual physician's performance to a relevant peer group can also support improvement. This performance comparison appeals to the competitiveness of physicians to improve and, to the extent the comparisons are public, provides even greater impetus to improve. Additionally, the timeliness of the performance assessment is important. Plans typically conduct their physician performance assessments at most annually, using data that is often at least a year old. Consequently, the assessments may preclude the timely detection of performance changes and yield results that are no longer valid. This can be particularly frustrating for physicians who are engaged in the process and making improvements.

There are a number of other ways plans can support physicians to improve their performance. For example, it would be useful to provide physicians guidance on the quality and costs of other providers to whom they refer. Although this is not a common practice among plans, without this information, physicians are likely to continue existing referral patterns even though the recipient providers may be

poor performers. Finally, providing peer-learning and support opportunities to physicians, including disseminating best practices and encouraging broader adoption would also be beneficial to help physicians improve performance.

Rewards for Good Performance. Key to any successful performance improvement effort is the inclusion of meaningful incentives that motivate and reward good results. However, the incentives currently associated with most health plans' physician performance measurement programs are minimal at best and do little to gain physicians' attention, engage them in the process or motivate them to improve. When incentives are offered, they are generally too inconsequential to be effective. Often, the incentive is limited to a physician designated as high performing in a plan's provider directory. Plans have been largely unwilling to offer rewards for better performance, in part because they do not want to increase aggregate payments, but also because they do not have a strong basis for penalizing low-performing physicians.

Significant Potential or a Potential Lost Opportunity?

Physician performance measurement, if credible and with relevant and robust improvement opportunities and rewards, offers tremendous promise to improve the quality and efficiency of care. Yet, most existing health plan programs have yielded information of limited value and usefulness, particularly for physicians and consumers. Many of these efforts also have been mired in underlying skepticism and distrust about plan motivations and methodological concerns, aggravated by sometimes conflicting results when plans pursue their own individual measurement efforts.

The measurement of physician performance has implications far beyond any single plan and its enrollees, extending to the population as a whole. Consequently, it is ill-advised to think of physician performance measurement as something other than a public good. Transcending competitive dynamics and encompassing a broader scope can position physician

performance measurement as a legitimate and valuable activity that yields demonstrable improvements in quality and costs. But, there are several critical challenges to moving physician performance measurement forward.

One challenge is that there is little consensus on what standards should guide these programs although the recent efforts in New York that require compliance to a prescribed set of standards may offer a good starting point toward national standardization. Additionally, the Consumer-Purchaser Disclosure Project, a consortium of purchasers, is working collaboratively with plans and providers to create a national set of principles to guide how health plans measure physicians' performance and report the information to consumers.¹³

Another important challenge is measuring physicians' performance across their entire patient panel, not piecemeal as is the case now with individual plans focusing only on their respective subset of patients. However, this requires combining data from all payers, including Medicare and Medicaid, to conduct an accurate assessment and likely will require some legal authority, such as the federal government, to mandate that it happen. The Centers for Medicare and Medicaid Services' (CMS) Generating Medicare Physician Quality Performance Measurement Results (GEM) project, which provides physician group practice performance data on a limited set of quality measures derived from Medicare Part B claims data offers a potential model and framework. The intended purpose of the project is to make this information available to Chartered Value Exchanges—regional collaboratives focused on health care quality and efficiency—to combine with commercial payer data to obtain a more complete profile of physician group practices.¹⁴

The absence of a convening entity with the necessary capacity, wherewithal and clout to neutralize existing competitive dynamics and champion physician performance measurement is another key challenge. In the current context, CMS may be the only candidate that fits this bill. A convener would be instrumental, if not essential, in helping to standardize the process and could serve as a central data repository

into which payers reported, improving efficiencies and eliminating conflicting results that ensue from individual plan efforts.

Finally, effective support for physicians willing to improve and robust rewards for physicians demonstrating good results are important. Otherwise, as the experience to date suggests, it is difficult, if not impossible, to engage physicians in the process. The support and rewards have to be of value to physicians to avoid distraction by competing demands. Although the challenges outlined here are formidable, failure to take the appropriate steps to improve the current state of physician performance measurement may result in a lost opportunity to improve the quality and efficiency of the underperforming U.S. healthcare system.

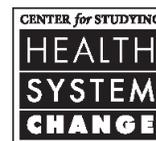
NOTES

1. Centers for Medicare and Medicaid Services (CMS), *National Health Expenditures: 2007 Highlights*, Washington, D.C.
2. Kaiser Family Foundation, *Health Care Spending in the United States and OECD Countries (Snapshots: Health Care Costs)*, Washington, D.C. (January 2007).
3. CMS, *National Health Expenditures: 2007 Highlights*. At a minimum, physicians influence spending for hospital care (32 percent of total national health expenditures), physician/clinical services (22 percent), and prescription drugs (10 percent).
4. Tynan, Ann, Allison Liebhaber and Paul B. Ginsburg, *A Health Plan Work in Progress: Hospital-Physician Price and Quality Transparency*, Research Brief No. 7, Center for Studying Health System Change, Washington, D.C. (August 2008).
5. Christianson, Jon, Paul B. Ginsburg and Debra A. Draper, "The Transition from Managed Care to Consumerism: A Community-Level Status Report," *Health Affairs*, Vol. 27, No. 5 (September/October 2008).
6. Draper, Debra A., Allison Liebhaber and Paul B. Ginsburg, *High-Performance Health Plan Networks: Early Experiences*, Issue Brief No. 111, Center for Studying Health System Change, Washington, D.C. (May 2007).
7. Pham, Hoangmai H., et al., "Redesigning Care Delivery in Response to a High-Performance Network: The Virginia Mason Medical

Center," *Health Affairs*, Vol. 26, No. 4 (July/August 2007).

8. "Regence BlueShield Sued Over 'Select' Network," *Seattle Post-Intelligencer* (Sept. 21, 2006).
9. Langston, Edward L., "Pay for Performance Programs," Report of the Board of Trustees, American Medical Association (December 2007).
10. "NCQA posts reviews of health plans under 2007 physician ranking settlement," *Health Plan & Provider Report* (Jan. 7, 2009).
11. Rosenthal, Meredith B., et al., "Paying for Quality: Providers' Incentives for Quality Improvement," *Health Affairs*, Vol. 23, No. 2 (March/April 2004); Rosenthal, Meredith B., and Richard G. Frank, "What is the Empirical Basis for Paying for Quality in Health Care?" *Medical Care Research and Review*, Vol. 64, No. 2 (April 2006).
12. Pham, et al. (July/August 2007).
13. Consumer-Purchaser Disclosure Project, <http://www.healthcaredisclosure.org/> (accessed March 3, 2009).
14. CMS, Generating Medicare Physician Quality Performance Measurement Results (GEM), <http://www.cms.hhs.gov/GEM/> (accessed March 3, 2009).

Funding Acknowledgement: This work was supported by the Robert Wood Johnson Foundation.



COMMENTARIES are published by the Center for Studying Health System Change.

600 Maryland Avenue, SW, Suite 550
Washington, DC 20024-2512
Tel: (202) 484-5261
Fax: (202) 484-9258
www.hschange.org

President: Paul B. Ginsburg