January 2009

The Law and Policy of Health Care Quality Reporting

Kristin Madison

Follow this and additional works at: http://scholarship.law.campbell.edu/clr

Part of the Health Law and Policy Commons

Recommended Citation

This Article is brought to you for free and open access by Scholarly Repository @ Campbell University School of Law. It has been accepted for inclusion in Campbell Law Review by an authorized administrator of Scholarly Repository @ Campbell University School of Law.
The Law and Policy of Health Care Quality Reporting

KRISTIN MADISON*

INTRODUCTION

Health care quality report cards have proliferated over the last decade. There are report cards for hospitals, physicians, medical groups, nursing homes, and health plans. There are report cards published by the federal government, state governments, state hospital associations, local and national nonprofit organizations, health insurers, and other commercial entities. Report cards vary significantly in the information they provide; quality measures may be based on structural characteristics of providers, processes of care delivery, patient outcomes, or patient experiences. Some report cards use stars, some use bars, and others simply publish quality-related statistics.

Each new report card offers the potential to improve the quality of medical care. Quality measures can help providers identify potential quality deficiencies, a crucial early step in the quality improvement process. Public and private regulators can use quality measures to identify providers giving low-quality care, and then take action to improve the quality of care delivered. Public report cards can also reinforce the professional imperative to improve quality, motivating providers through their reputational impact and influence on the competitive process. Providers seeking to attract patients directly or through referrals may become more successful if they can establish a higher level of quality. Ultimately, report cards can correct information failures that plague health care markets and help patients obtain the care they desire.

At the same time, however, each new report card brings with it a potential for harm. Just as report cards can change provider behavior in beneficial ways, they can change provider behavior in problematic ways. Physicians concerned about their performance may try to obtain higher scores by selecting as patients people that they think will help them perform well. If report cards include outcome-based measures, for example, they may turn away sicker patients. Providers may begin to “teach to the test,” focusing on improving performance in areas that

* Professor of Law, University of Pennsylvania Law School, and Senior Fellow, Leonard Davis Institute of Health Economics.

215

Published by Scholarly Repository @ Campbell University School of Law, 2009
are measured, while neglecting or failing to improve care in other areas, even if the other areas are equally important. But perhaps the most significant potential for harm arises when report cards are inaccurate or misleading. If providers unknowingly rely on inaccurate quality measures, they may do nothing when they might otherwise have looked into a quality problem, or they may waste resources or even worsen care in an effort to address a problem that does not exist. If payers or patients rely on poor-quality measures, they may select or reward physicians and hospitals that provide low-quality care, while moving away from providers that offer high-quality care. Misinformation may mean worse care.

Will report cards improve or worsen quality? Or might they have no effect at all? While evidence relevant to these questions is beginning to accumulate, definitive answers remain elusive, at least for now. Ultimately, report cards’ benefits and costs will depend on the ways that quality measurement and reporting evolve in the next few years. Numerous stakeholders will shape this evolutionary process, including providers, government agencies, regulators, insurers, employers, and consumers, as well as report card producers themselves. By addressing report cards’ shortcomings while enhancing their benefits, these groups can increase the likelihood that quality reporting promotes the growth of high-quality health care in the United States.

Part I of this Essay provides an overview of hospital and physician quality reporting in the United States today. It compiles examples of quality report cards, examines empirical evidence of report cards’ impact, and discusses report cards’ problems, including barriers to consumer use and deficiencies in content and quality. Part II reviews legal issues that have arisen with respect to quality reporting, focusing on controversies surrounding health insurers’ use and publication of physician performance measures. It traces the roles of various stakeholders in responding to these controversies and shows how the resolutions reached address broader policy concerns about quality measures. Drawing on the discussion in Parts I and II, Part III describes how various stakeholders can take fuller advantage of quality reporting’s benefits.

I. HEALTH CARE QUALITY REPORTING

Health care quality reporting is not a new phenomenon. In the early 1970s, for example, the Pennsylvania Commissioner of Public Insurance published a one-page “shopper’s guide” to Philadelphia hos-
HEALTH CARE QUALITY REPORTING

pitals. In the 1980s and early 1990s, several states and the predecessor to the current Centers for Medicare and Medicaid Services (CMS) began to publish provider-specific health care quality statistics. Recently, however, the quality reporting trend has greatly accelerated. While many factors have undoubtedly contributed to this acceleration, two stand out: increasing attention to health care quality concerns and declining information technology costs. Two widely cited Institute of Medicine reports, To Err Is Human and Crossing the Quality Chasm, raised awareness about the health care system's quality shortcomings among medical professionals, policymakers, and, through media coverage, the general public. Subsequent studies have continued to document problems in health care provision. Concerns about deficits in the quality of care, whatever their source, help justify the creation of reporting mechanisms. At the same time, the declining costs of information technology make it easier and cheaper to collect, analyze, publish, and access quality-related information. The growth of the Internet has made possible widespread distribution of quality report cards.

A. The Nature of Quality Report Cards

While quality report cards exist for many health-related entities, including nursing homes and health plans, this Essay will focus on


2. See id.; see also Timothy Stoltzfus Jost, Oversight of the Quality of Medical Care: Regulation, Management, or the Market?, 37 Ariz. L. Rev. 825, 837 (1995) (describing early quality reporting initiatives).

3. For a more thorough discussion of the many trends underlying the development of consumer-oriented quality reporting, see generally Tomes, supra note 1.


7. For more discussion of information technology's impact on health care quality reporting and on quality regulation more generally, see Kristin Madison, Health Care Quality Regulation in an Information Age, 40 U.C. Davis L. Rev. 1577, 1595-1601 (2007); see also Tomes, supra note 1 (describing effects of information technology).

report cards for two types of providers: hospitals and physicians. Within each of these categories, report cards take many forms. Turning first to hospital report cards, the entities that choose to publish quality-related information about hospitals are quite diverse. Producers of report cards include the United States Department of Health and Human Services (HHS);¹⁰ states such as Florida, New York, and Pennsylvania;¹¹ national organizations such as the Joint Commission,¹² which accredits hospitals, and the Leapfrog Group,¹³ which promotes health care safety and quality; local and regional organizations, such as the North Carolina Center for Hospital Quality and Patient Safety;¹⁴ commercial entities such as HealthGrades;¹⁵ and health plans.¹⁶

The nature of measures included in these report cards varies considerably. Process of care measures are common among hospital report card websites. The HHS Hospital Compare website, for example, includes quality indicators based on adherence to care guidelines, such as the "[p]ercent of surgery patients who were given an antibiotic at the right time (within one hour before surgery) to help prevent infection," and the "[p]ercent of [h]eart [a]ttack [p]atients [g]iven [b]eta [b]locker at [d]ischarge."¹⁷ Some sites report hospital structural characteristics that may be associated with quality; the Leapfrog Group

---

10. See Medicare.gov, Hospital Compare (Feb. 11, 2009), http://www.hospitalcompare.hhs.gov.
site, for example, reports on whether hospitals have adopted computerized prescriber order entry systems. Others report on problems that arise during care such as infections or surgery on the wrong body part. A few sites, such as Florida’s, report on patient outcomes such as mortality following bypass surgery or hip replacement. Hospital report cards may also include information on patient experiences. The Hospital Compare site publishes measures drawn from the Hospital Consumer Assessment of Healthcare Providers and Systems survey, such as measures based on communication with hospital nurses, pain management, cleanliness, and quietness.

Physician quality reporting is less common than hospital quality reporting, probably because of the difficulty of developing and collecting meaningful statistical quality measures when each physician sees relatively few patients. Perhaps the most common type of physician reporting is that available in state licensure databases, which are not really report cards, but instead listings of licensed physicians along with information about characteristics potentially associated with quality, such as whether the physician is board-certified, has been subject to professional discipline, or has paid malpractice claims. The scope of information in such databases varies considerably from state to state. Commercial entities such as HealthGrades also provide such information. States such as New Jersey, New York, and Pennsylvania calculate a limited set of physician-specific quality measures, such as cardiac bypass surgery mortality rates and hip and knee

replacement infection rates.\textsuperscript{26} Health insurers may provide information related to physician quality by tiering physicians according to performance or creating high-performance networks.\textsuperscript{27} Medicare began tracking physician quality measures on a voluntary basis—about sixteen percent of physicians participated in the program in 2007—but has not published this data.\textsuperscript{28} Consumers offer their own assessments of individual physician quality on public websites.\textsuperscript{29} In some communities, quality ratings are also available for physician groups.\textsuperscript{30}

B. The Impact of Quality Report Cards

Part I.A's compilation of examples of quality report cards suggests that quality reporting has become a widespread phenomenon. Health care provider report cards have indeed become quite common. Their ubiquity, however, does not necessarily translate into significant impacts on health care quality. First, the fact that report cards exist does not mean that people regularly read them. Second, even if patients and providers are familiar with report cards, they may not respond to their contents by altering their behavior in ways that

\begin{itemize}
  \item \textsuperscript{29} See, e.g., RateMDs.com, Doctor Ratings and Reviews, \url{http://www.ratemds.com} (last visited Mar. 7, 2009) (providing physician ratings based on survey of website users).
\end{itemize}
improve health care quality. The real-life impact of quality report cards must be assessed empirically.

1. Patients' Report Card Use

The pervasiveness of "best physicians" issues of local magazines hints at widespread public interest in health care provider ratings. More systematic evidence confirms such interest.\(^{31}\) At the same time, however, surveys show that only a minority of the public has seen report cards, and only a subset of those who have seen them actually use them. The majority of respondents to a 1996 telephone survey of Pennsylvania bypass surgery patients said that they were interested in quality ratings and that they would change surgeons if they discovered that their surgeon had a higher than expected mortality rate.\(^{32}\) But only twelve percent of the patients knew about the availability of bypass surgery report cards, and an even smaller percentage made a decision based on them.\(^{33}\) The low rates of patient awareness should obviously concern report card advocates, given the relevance of the report card results to the patients surveyed. However, the extent to which early studies such as this one retain their relevance is open to question. The survey predated the growth of the Internet, through which most report cards are available today. The relative ease of locating and using Internet-based report card information would seem to greatly expand access to report cards.

Recent surveys suggest that access to quality reporting has in fact expanded. In 2008, for example, one survey found that about twenty percent of Americans saw comparative quality information on hospitals.\(^{34}\) The same survey found that about twelve percent of respondents saw information on physicians, a result consistent with the lower availability of physician report cards.\(^{35}\) Just as for the 1996 Penn-


\(^{33}\) Id.

\(^{34}\) KAISER FAMILY FOUND., 2008 UPDATE ON CONSUMERS' VIEWS OF PATIENT SAFETY AND QUALITY INFORMATION 6 (2008), available at http://www.kff.org/kaiserpolls/upload/7819.pdf. Only six percent of respondents reported having heard of the Department of Health and Human Services' hospital comparison website. Id. at 8.

\(^{35}\) Id. at 6. Use of quality ratings may be higher for certain groups. For example, about a third of respondents to a 2008 survey of Internet users between the ages of
sylvania study, however, the percentage of respondents who actually used report card data was lower than these numbers imply—about seven percent for hospitals, and about six percent for physicians. 36 Similarly, a survey of California Internet users found that the percentage who saw physician ratings reached twenty-two percent in 2007, but that only five percent of respondents considered a change of physicians based on the ratings, and only two percent actually did so. 37

Many commentators label these usage levels as “low,” and indeed they are in absolute terms. Most Americans do not use health care quality report cards. The statistics could be interpreted in another way, however; these numbers could be viewed as quite substantial, given that many people have little reason to look over health care quality ratings, much less search them out. Healthy individuals may not read report cards because they have little reason to visit health care providers. Individuals who do use hospital or physician services may not bother consulting report cards if their health plans or geographical circumstances constrain their choice of providers. And individuals who are satisfied with their current providers may ignore report cards, despite their lack of knowledge about their current providers’ clinical quality. Patients who see no reason to switch providers would be unlikely to devote much time to report cards. In 2007, only about eight percent of Americans chose a new primary care practitioner, and about seven percent chose a new specialist. 38 These statistics offer one reason why relatively few people are seeking out report cards. 39

---


36. KAISER FAMILY FOUND., supra note 34, at 7.


39. Of course, for report card advocates, some of these explanations for low report card use may be more reassuring than others. If report cards are to realize their full potential for increasing quality, even satisfied patients should seek out report cards on providers’ clinical quality. Poor report card ratings might influence previously satisfied patients’ choices of providers. The more patients who use report cards, the greater the likelihood that report cards will influence behavior.
On the other hand, surveys suggest that report cards lag behind other sources of information for patients selecting new physicians. For example, one survey found that while about eleven percent of patients used the Internet in selecting a primary care practitioner, close to twenty-seven percent of patients relied only on friends and relatives. 40 Similarly, patients were more likely to rely on physician referrals, friends or relatives, or health plans in selecting specialists than to rely on the Internet. 41 Ultimately, survey evidence indicates that patients are using report cards to make health care decisions, but in limited numbers.

2. Effects of Report Cards on Market Share

Survey evidence offers only one perspective on the impact of report cards, and it has the limitation that it can show only what respondents say they do. An alternative approach to evaluating report cards' effects is to use objective empirical evidence to assess what people actually do in response to report cards. This approach has the advantage of taking into consideration not just patients' conscious decisions based on their own examinations of report cards, but also any other pathways through which report cards might affect behavior. For example, patients' physicians, friends, neighbors, or even their insurers might have seen report card results and then taken steps to steer patients toward or away from particular providers. Typically, studies taking a non-survey based empirical approach examine whether providers' treatment volumes or market shares change in response to report card results.

Some of these studies find effects, while others do not. For example, a study from the early 1990s found that hospitals and physicians that performed better on the New York cardiac surgery report card gained market share. 42 A 2004 study determined that hospitals that performed poorly on the same report card lost relatively healthy patients to other hospitals. 43 Another 2004 study of several report cards in New York and California concluded that for hospitals

40. Tu & Lauer, supra note 38, at 4.
41. Id.
42. Dana B. Mukamel & Alvin I. Mushlin, Quality of Care Information Makes a Difference: An Analysis of Market Share and Price Changes After Publication of the New York State Cardiac Surgery Mortality Reports, 36 MED. CARE 945 (1998).
reported to be either high-quality or low-quality outliers, "[v]olume effects were modest, transient, and largely limited to white Medicare patients in New York." 44 A 2006 study found that the New York report cards did not affect market share. 45 Citing a range of conflicting studies published through early 2006, the most recent systematic review of studies concluded that, overall, report cards did not alter choices of treatment hospitals. 46

More recent studies, however, suggest that under the right conditions, report cards may alter patient choices. A 2008 study explained that report cards will not shift market share if they do not inject into the market any new information about quality; market participants may already have views about the quality of providers that are simply reinforced by report cards. 47 Examining the period when New York cardiac report cards were first introduced, the authors showed that when report cards did contain negative new information about hospitals, these hospitals subsequently lost market share. 48 Interestingly, however, they did not show a systematic shift of patients toward hospitals with positive news. 49

While many studies of market share changes focus on cardiac report cards, a 2008 working paper examined the impact of report cards in a different setting: fertility clinics. 50 Report cards would seem to have great potential in this setting. As the authors explain, birth rate measures are easy to understand, facilitating report card use; fertility

44. Patrick S. Romano & Hong Zhou, Do Well-Publicized Risk-Adjusted Outcomes Reports Affect Hospital Volume?, 42 MED. CARE 367 (2004). Other studies have also found that report card effects may vary according to patient characteristics. For example, a study of the effect of report cards on kidney transplant centers found "some evidence that report cards influence younger and college-educated patients," but that "overall report cards do not affect demand." David H. Howard & Bruce Kaplan, Do Report Cards Influence Hospital Choice? The Case of Kidney Transplantation, 43 INQUIRY 150, 150 (2006).


48. Id.

49. Id. at 1207.

services are not emergency services, giving patients time to seek out the provider-specific information available in report cards; and fertility patients are disproportionately young and highly educated, increasing the likelihood that they would have the appropriate knowledge and skills to take advantage of report cards. The study found that after fertility clinics began to publicly report birth rates through a national report card system, clinics with higher birth rates gained market share. The study suggests that report cards can in some circumstances influence a patient's choice of providers.

3. **Effects of Report Cards on Quality**

Provider selection is a potentially important mechanism by which report cards may improve quality, but it is not the only one. Even if few patients actually select alternative providers, health care providers may respond to report cards by trying to improve their own quality. Providers' concern about their professional reputation among peers or the public, or more simply a desire to deliver the highest quality of care possible, may motivate quality improvement. One way to examine the aggregate effects of report cards' various quality improvement mechanisms is to look at whether there is a connection between report card publication and quality improvement. Evidence of a connection could take the form of quality improvement activities undertaken in response to report cards or, alternatively, actual post-report card quality improvement as indicated by the quality measures themselves.

In the quality effect context, just as in the market share context, the most frequently studied report card is probably the New York bypass surgery report card. One article offered detailed descriptions of the actions of specific New York hospitals in light of their report card performance. While it acknowledged that there was little evidence that mediocre hospitals sought to improve their performance in response to their report card results, it found that hospitals that performed especially poorly implemented many changes in their surgery programs in order to improve quality. Consistent with this anecdotal evidence of quality improvement activities, a study using patient data to evaluate the New York report card's effects concluded that poorly

51. *Id.* at 19-21.
52. *Id.* at 19.
53. See, e.g., Jha & Epstein, *supra* note 45.
55. See *id.* at 48.
56. *Id.* at 42-45.
performing hospitals experienced a decline in mortality results soon after the publication of the report.\textsuperscript{57} In addition, a 2006 study revealed that the surgeons with the highest mortality rates on these report cards were more likely than others to discontinue their practices after the report card release.\textsuperscript{58}

Studies of other types of report cards also highlight public reporting’s potential to make a difference in health care quality. An assessment of Wisconsin hospital report cards, for example, found that engagement in quality improvement activities was higher among hospitals subject to public reporting, relative to hospitals that received confidential report cards or no report cards at all.\textsuperscript{59} A recent systematic review concluded that evidence from eleven hospital-level studies suggested that report card publication spurred hospital quality improvement activities.\textsuperscript{60}

\textbf{C. The Problems of Quality Report Cards}

Subpart B concluded that while patients are beginning to use report cards, such patients remain in the minority; that the evidence that report cards redirect patients to higher-quality providers is at best mixed, although some of the most recent studies show an effect; and that reporting has likely spurred hospital quality improvement activities, although studies of the link between reporting and measurable quality improvement are limited. These conclusions imply that while report cards have the potential to improve quality and probably have begun to do so, they still have a long way to go to reach their full potential for quality improvement. This is an optimistic view. Report card skeptics would argue that they are unlikely ever to reach their full potential, while report card critics would argue that they might even make things worse. At the core of all three of these views lie concerns about the problems of quality report cards. This subpart describes these problems.

\textsuperscript{57} \textit{See} Cutler et al., \textit{supra} note 43, at 345.

\textsuperscript{58} Jha & Epstein, \textit{supra} note 45, at 844.

\textsuperscript{59} Judith H. Hibbard et al., \textit{Does Publicizing Hospital Performance Stimulate Quality Improvement Efforts?}, 22 \textit{Health Aff.} 84, 90-91 (2003). Note, however, that hospitals were not randomly assigned to the publicly-reported category, raising questions about whether they may differ in systematic ways from the other hospitals. \textit{See id.} at 87, 92-93.

\textsuperscript{60} Fung et al., \textit{supra} note 46, at 111. There is little systematic evidence on the impact of quality reporting on physician services. \textit{Id.}
1. Barriers to Report Card Use

There are many potential explanations for why report card use is not yet widespread. Subpart B.1 described one possibility: many consumer-patients may not see the need for report cards, because they do not need a provider, have a limited choice of providers, or are satisfied with their current providers. Survey evidence suggests that many people are not aware of existing quality differentials, which would dampen any enthusiasm they might otherwise have for searching out comparative quality information. Another possibility is that report cards are relatively new, and it may simply take time for people to become aware of their existence. Finding them on the Internet may take even longer; searching for the term “hospital quality” or “hospital report card” does not always turn up relevant report cards, even if they do exist. In addition, some people face considerable barriers to report card use, such as a lack of Internet access or cognitive or physical limitations on the ability to acquire and use information. Hard-to-understand formatting can significantly magnify such barriers, further discouraging report card use, as well as leading to improper use.

2. The Content of Report Cards

Other explanations for limited use of report cards stem from their content. Simply put, people will not use report cards unless they perceive their content to be “useful.” For some users, report cards fall short along this dimension. One fact readily apparent from Subpart A’s survey of the report card landscape is that although report cards are increasingly common, relatively few focus on physician-specific quality measures. Hospital quality measures can of course be useful; poor-quality hospital care often has serious health consequences. For many people, however, hospital admissions are rare, and those that do occur are guided by referring physicians, heavily influenced by geographical considerations, or constrained by health plans. By contrast,

61. See Kaiser Family Found., supra note 34, at 9 (reporting that only forty-one percent of respondents perceived “big differences” in local hospital quality, thirty-three percent in specialist quality, and thirty percent in primary care physician quality); see also Harris & Buntin, supra note 31, at 10 (describing reasons consumers decline to use hospital quality information, including “[h]igh levels of satisfaction with one’s own provider and lack of perceived quality differences across providers”).

patients visit physicians frequently and often select physicians (especially primary care practitioners) with little professional help. Also, patients generally have a broader choice of physicians than of hospitals. The relative lack of report cards on physician services is therefore highly problematic.

Another limitation on the usefulness of report cards is the nature of their performance measures. Developing appropriate measures of clinical quality can be difficult. The ideal measures, whether structure, process, or outcome-oriented, would be based on reliable evidence that establishes an association with dimensions of quality valued by patients. However, there are many areas in health care for which systematic evidence of the best treatment approaches is lacking. Furthermore, particularly for process-oriented and outcome-oriented measures, limitations on available data and statistical techniques complicate efforts to develop statistically reliable quality measures that differentiate providers sufficiently to make quality ratings meaningful. And collecting the right data for report cards can be burdensome.

The challenges inherent in developing report cards have meant that even the most developed report cards contain limited data. The ideal report card for a particular patient at a particular time would be tailored to that patient's needs. First, clinical quality measures would be for the condition the patient has or the treatment the patient needs. A patient in need of obstetrical services would want to know about the quality of a physician's obstetrical services, as well as the quality of the obstetrics unit within the hospital. A patient in need of cancer treatment would be interested in how providers' cancer patients fare, particularly patients with similar types of cancer, at similar stages. There are a few report cards that contain measures directed at a particular type of patient; the bypass surgery report cards are one example. But what if I am an obstetrics patient? Is a hospital's overall infection rate a key measure to consider in choosing a hospital? Or might infection rates vary across units within the hospital? Would the Hospital Compare clinical quality statistics on cardiac care, pneumonia care, and certain infection prevention measures be relevant? What if I need surgery, but not cardiac surgery or knee surgery? What report cards would be helpful to my choice of providers? Many existing quality measures are very broad, potentially masking variations within an institution. The overall hospital surgical site infection rate may be one example. Others are focused on just a few conditions or treatments, leaving patients with other conditions and patients in need of other

63. See Medicare.gov, supra note 10.
treatments to wonder about the report cards’ relevance. The more that an individual provider’s quality is correlated across treatment types and settings, the less of a problem it is to focus on just a few conditions. But without more information about such a correlation, potential report card users cannot simply assume that report cards focusing on different conditions or treatments provide useful information.

Second, in an ideal world, report card measures would focus on clinical outcomes. Patients care about outcome measures such as lower mortality rates, not whether physicians are board-certified or the frequency of beta blocker prescriptions in hospital settings. These alternative measures are imperfect at best, and are used mainly because it can be costly and difficult to measure outcomes and to risk-adjust them properly, so that they reflect the providers’ quality, rather than underlying patient characteristics. In practice, very few report cards measure outcomes, and the outcomes they capture tend to be a limited subset of what patients may care about. The Hospital Compare report card, for example, offers three outcome measures: heart attack mortality rates, heart failure mortality rates, and pneumonia mortality rates. Pennsylvania’s knee and hip replacement report cards report on deep joint infections and device problems, blood clots, wound infections, re-admission rates, and post-operative lengths of stay. But other types of patients may wonder about other mortality rates, and all patients may wonder about other kinds of outcome measures. What are the mortality rates for neurosurgery patients? To what extent is a provider’s treatment successful in restoring functioning? How long is the typical recovery period? There are many variations on these types of questions, and few of them are addressed in existing report cards.

While report cards’ objective clinical quality measures may fall short of the ideal, many patients would argue that their improvement would only go part of the way toward making quality report cards truly useful. Patients care about more than just objective clinical quality measures; they care a great deal about non-clinical quality measures such as convenience and amenities. A 2007 survey found that the share of patients who considered convenience a factor in choosing a physician was on par with the share who considered reputation or perceived quality a factor. A 2008 working paper found that while lower pneumonia mortality rates contributed to higher patient demand for a hospital’s services, increases in perceived hospital amenities were asso-

64. Id.
65. PA. HEALTH CARE COST CONTAINMENT COUNCIL, supra note 26, at 7-11.
66. Tu & Lauer, supra note 38, at 6.
associated with much greater increases in demand.\textsuperscript{67} One could easily imagine that factors such as convenient parking, better hospital food, or shorter appointment waiting times might be relevant to patients.

Patients also highly value information about other patients' experiences with physicians and hospitals.\textsuperscript{68} The Institute of Medicine has included among its aims for quality improvement "patient-centered" care; systematic surveys of patient experiences support this goal.\textsuperscript{69} Several websites have begun to publish results from such surveys. For example, in 2007, Massachusetts Health Quality Partners published doctor's office-specific ratings based on surveys returned by 51,000 adult patients about factors such as how well doctors communicate with patients, how well doctors coordinate patient care, and the helpfulness of office staff.\textsuperscript{70} And in 2008, the Hospital Compare website introduced a set of quality measures drawn from the Hospital Consumer Assessment of Healthcare Providers and Systems, a survey which asks discharged hospital patients for ratings in areas such as communication with nurses and doctors, pain management, and the provision of discharge information, as well as for ratings of the hospital as a whole.\textsuperscript{71} Many of the factors assessed by these surveys are clinically relevant, and studies have found a positive (although imperfect) correlation between patient experience and other kinds of clinical quality measures.\textsuperscript{72}

\textsuperscript{67} Dana Goldman & John A. Romley, \textit{Hospitals as Hotels: The Role of Patient Amenities in Hospital Demand} (Nat'l Bureau Econ. Research, Working Paper No. 14619, 2008) (noting that "[a]menities such as good food, attentive staff, and pleasant surroundings may play an important role in hospital demand" and finding that a "one-standard-deviation increase in amenities raises a hospital's demand by 38.4\% on average, whereas demand is substantially less responsive to clinical quality as measured by pneumonia mortality.").

\textsuperscript{68} See \textit{HARRIS & BUNTIN}, supra note 31, at 3 (citing sources); see also Beckey Bright, \textit{Patient Surveys Seen as Reliable for Rating Health-Care Providers}, \textit{WALL ST. J.}, Feb. 23, 2008 (describing poll result that ninety-one percent of respondents would be likely to consult health plan enrollees' ratings on trust, communications, medical knowledge, availability, and office environment when selecting a physician). Individual patients also frequently share their experiences with health care providers through the Internet, yet another example of widespread interest in patient experiences.

\textsuperscript{69} See \textit{BD. ON HEALTH CARE SERVS., INST. OF MED., PERFORMANCE MEASUREMENT: ACCELERATING IMPROVEMENT} 118-19 (2006) (describing quality aims, including patient centeredness, and providing examples of patient-centered performance measures).


\textsuperscript{71} See \textit{HCAHPS}, supra note 22, at 1.

\textsuperscript{72} See \textit{HARRIS & BUNTIN}, supra note 31, at 7, and sources cited therein.
Given patients’ interest in aspects of provider performance other than those captured by technical clinical quality measures, one potential way to increase the impact of report cards is to expand the effort devoted to reporting on factors such as amenities and patient experiences. Perhaps consumer preferences ought to make a difference in the health care context, just as they do in other contexts. Report card-induced selection and competition based on these characteristics may very well benefit patients.

At the same time, though, there is reason to be cautious in expanding report cards’ focus on patient experiences, especially if it comes at the cost of directing attention and resources away from the development of other quality measures. One reason to temper enthusiasm about patient experience measures involves the economic principle of comparative advantage.73 Under this principle, even if one entity is better than another at producing each of two goods, it is economically more efficient for the first entity to produce the good that it has the bigger advantage in producing, while the second entity produces the other good.74 This reasoning suggests that even if report card providers are superior to consumers at collecting both experience measures and technical quality measures, it may make sense for report card providers to increase their focus on technical quality measures. After all, consumers can collect some information about patient experiences. Prospective patients can ask friends or family about their experiences with specific providers, an approach that is particularly beneficial if the prospective patients know that these family and friends share their preferences. Friends and family are in practice a very important source of information about quality.75 By contrast, consumers would find it nearly impossible to assemble technical clinical quality data on their own.

This is not to say that systematic patient experience surveys are not useful; they certainly are. Some people may not know others they can ask or may be reluctant to ask questions that are too personal. They may not have the time to talk to a sufficient number of patients to develop a meaningful assessment of patient experience with a specific

73. See John B. Taylor & Akila Weerapana, Principles of Macroeconomics 6 (6th ed. 2009) (“In general, a person or group of people has a comparative advantage in producing one good relative to another good if that person or group can produce that good with comparatively less time, effort, or resources than another person or group can produce that good.”).

74. See id. at 6 (defining comparative advantage); see also id. at ch. 29 (discussing comparative advantage in more detail).

75. See, e.g., Harris & Buntin, supra note 31, at 9 (describing patterns of use of various types of quality information).
provider. They may not ask the right questions. Well-designed surveys can capture elements of patient experience that prospective patients may benefit from learning about, but simply not think to ask. In addition, experience surveys provide feedback to providers, who can then look for ways to improve satisfaction with their services. But, ultimately, the question is not just whether patient experience surveys are beneficial; the question is what level of resources to devote to them when these resources could instead support clinical quality measures that consumers have no way of collecting on their own.

If you asked patients whether they would prefer to see patient experience measures or clinical quality measures, they might very well say patient experience measures. People sometimes discount clinical quality measures. For example, one survey found that the percentage of respondents who would prefer a surgeon who has treated friends or family nearly equaled the percentage who would choose a surgeon whose quality had been rated higher.76 Similarly, the survey found that the percentage of respondents who would choose a familiar hospital (fifty-nine percent) greatly exceeded the percentage who would choose a more highly rated hospital (thirty-five percent).77 And in a direct comparison of patient experience measures to more technical types of quality measures, an experimental study of primary care provider report cards found that nearly one-third of participants chose providers rated highly on interpersonal quality (in areas such as communication, courtesy, and promptness) but poorly in the delivery of health care services.78

If the preference for interpersonal quality as measured by patient experience surveys is strong, even on occasions when it is accompanied by subpar delivery of care or poor health care outcomes, then perhaps report cards' focus on patient experiences should be further intensified. However, the concern is that patients making decisions based primarily on information from friends, family, or patient experience surveys, might be doing so because they do not understand clinical quality measures or their importance.79 If patient decisions are driven by a lack of understanding, rather than preferences, then

76. KAISER FAMILY FOUND., supra note 34, at 10.
77. Id.
78. See Constance H. Fung et al., Patients' Preferences for Technical Versus Interpersonal Quality When Selecting a Primary Care Physician, 40 HEALTH SERVS. RES. 957, 971 (2005).
79. See HARRIS & BUNTIN, supra note 31, at 3 ("There is some indication that disinterest in clinical performance measures found in focus group studies may be driven by consumers' perceived lack of expertise regarding the clinical issues underlying performance data.").
public benefit-minded report card designers should try to address the underlying comprehension problem.

3. The Quality of Report Cards

One reason that patients may be reluctant to rely on clinical quality measures in choosing providers is a concern that the measures do not accurately reflect quality. Health care providers, policymakers, researchers, and report card designers share this concern. As described in Subpart C.2, accurate quality measurement can be very difficult to achieve. High-quality quality measures are nonetheless critical. Just as well-designed health care quality measures can contribute to quality improvement, poorly designed quality measures can worsen quality. First, they can misdirect patients to poor-quality providers and away from high-quality providers. Patients who rely on problematic health care quality report cards will be worse off as a result. Second, they can misdirect quality improvement efforts. Quality ratings provide information to health care providers, managers, and payers, as well as to patients. Providers may focus on improving areas that are actually already high-quality, thus wasting resources. They might also overlook areas in need of quality improvement, thus sacrificing opportunities to improve patient treatment. Faulty quality measurements may lead third parties to reward poor-quality care or penalize high-quality care, thus skewing provider incentives.

Third, because providers may be sensitive to these impacts of poorly designed quality measures, they may alter their behavior accordingly, with potentially problematic consequences. If a provider suspects that an outcome measure is not adequately adjusted for the underlying health risks of patients, for example, the provider might turn away sicker patients that the provider would otherwise treat. Or a provider might choose to provide a procedure to a patient to whom they might otherwise recommend another form of treatment. Such behavior might increase health care disparities as well as worsen care overall. Many authors have suggested that physicians might avoid treating patients who are members of racial minorities, because they may be concerned that such patients would be less likely to adhere to treatment advice or more likely to have poor outcomes.

80. See, e.g., Rachel M. Werner & David A. Asch, The Unintended Consequences of Publicly Reporting Quality Information, 293 JAMA 1239, 1239 (2005) (reviewing potential benefits of report cards but also detailing how they “may have unintended and negative consequences on health care”).

81. See, e.g., Lawrence P. Casalino et al., Will Pay-for-Performance and Quality Reporting Affect Health Care Disparities?, 26 HEALTH AFF. w407, w408 (2007); Huw T.
Empirical studies have shown that the concern that providers might react to quality ratings by altering treatment patterns in problematic ways is legitimate. One study showed that bypass surgery report cards were associated with an increased disparity in bypass surgery use between white patients, on the one hand, and black and Hispanic patients, on the other.\footnote{Davies et al., Health Care Report Cards: Implications for Vulnerable Patient Groups and the Organizations Providing Them Care, 27 J. Health Pol. Pol'y & L. 379, 393 (2002); Werner & Asch, supra note 80, at 1243.} Another, broader study of Medicare data on cardiac patients from 1987 to 1994 suggested that during their first four years, New York and Pennsylvania bypass surgery report cards “led to substantial selection by providers,” as indicated by a decline in illness severity of patients receiving bypass surgery and declines in certain cardiac procedures for sicker heart attack patients.\footnote{Rachel M. Werner et al., Racial Profiling: The Unintended Consequences of Coronary Artery Bypass Graft Report Cards, 111 Circulation 1257 (2005).} It also found that the report cards led to “improved matching of patients with hospitals,” in the sense that more severely ill patients were increasingly treated at teaching hospitals.\footnote{David Dranove et al., Is More Information Better? The Effects of “Report Cards” on Health Care Providers, 111 J. Pol. Econ. 555, 583 (2003).} A report card-induced change in patient mix is not necessarily a bad thing; for example, if physicians were too aggressive pre-report card in providing treatment to severely ill patients, resulting in worse outcomes for these patients, then a finding of declining illness severity would be a positive result. Such changes are in fact one goal of report cards. Thus, to determine whether report card-induced changes are socially beneficial, it is important to look at their impact on health outcomes. The study authors found that in the aggregate, the changes “led to higher levels of resource use and to worse health outcomes, particularly for sicker patients,” leading the authors to conclude that “these report cards decreased social welfare.”\footnote{Id. at 556.}

All of these potential problems highlight the importance of high-quality report cards. At a minimum, performance measures should accurately reflect some dimension of a provider’s “true” underlying quality. This means that structure- and process-oriented quality measures should reflect practices known to result in improved health care outcomes, and that outcome measures should be properly risk-adjusted. The measures should be based on current, accurate data. Ideally, report cards would also be designed to capture the many dimensions of quality valued by patients, and to reduce the likelihood


84. Id. at 556.

85. Id.
that a single-minded focus on report card measures would degrade quality along other, equally important dimensions. Unintended consequences are always a concern.\textsuperscript{86}

It can be difficult to assess the quality of report cards, in part because it is not clear what the gold standard should be when every measurement methodology has a potential flaw. A few studies have raised concerns based on comparisons of quality report cards, however. A study examining the hospitals that \textit{U.S. News & World Report} identified as among “America’s Best Hospitals” for “Heart & Heart Surgery” found lower risk-adjusted mortality rates among the “Best” hospitals than among non-ranked hospitals, but also found that the list omitted many others that had low mortality rates.\textsuperscript{87} A study comparing five websites’ ratings for hospitals in the Boston area found that they failed to agree because they used different measures, considered the experiences of different patients, and assessed different reporting periods.\textsuperscript{88} It should not be surprising that different measurement methodologies and databases generate different rankings, but these results do serve as a reminder of the importance of careful scrutiny of report card content.

Other evidence supports a conclusion that report cards can be sufficiently informative to guide patient choice. The authors of a 2006 study of New York’s bypass surgery report cards, for example, concluded that “users who picked a top-performing hospital or surgeon from the latest available report had approximately half the chance of dying as those who picked a hospital or surgeon from the bottom quartile.”\textsuperscript{89} Another 2006 study found that the Hospital Compare process of care quality measures predicted differences in hospital risk-adjusted

\textsuperscript{86} For a discussion of the problems that have arisen in the context of measures of pneumonia-related quality of care, see Mark L. Metersky, \textit{Measuring the Performance of Performance Measurement}, 168 \textit{ARCHIVES INTERNAL MED.} 347 (2008).

\textsuperscript{87} Oliver J. Wang et al., “America’s Best Hospitals” in the Treatment of Acute Myocardial Infarction, 167 \textit{ARCHIVES INTERNAL MED.} 1345, 1345 (2007) (“On average, admission to a ranked hospital for AMI was associated with a lower risk of 30-day mortality, although about one-third of the ranked hospitals fell outside the best performing quartile based on RSMR [risk-standardized mortality rates]. Although ranked hospitals were much more likely to have an SMR significantly less than 1, many more non-ranked hospitals had this distinction.”).

\textsuperscript{88} Michael B. Rothberg et al., \textit{Choosing the Best Hospital: The Limitations of Public Quality Reporting}, 27 \textit{HEALTH AFF.} 1680, 1680 (2008). For another article comparing the characteristics of report cards, see Michael J. Leonardi et al., \textit{Publicly Available Hospital Comparison Web Sites}, 142 \textit{ARCHIVES OF SURGERY} 863 (2007).

\textsuperscript{89} Jha & Epstein, supra note 45, at 844.
mortality rates. Similarly, a 2007 study found that better performance on a set of commonly used process of care quality indicators was associated with lower mortality. The fact that these studies find a relationship between report card quality measures and quality outcomes is reassuring, but it does not guarantee that all report cards are carefully constructed.

II. A Case Study: Health Plan Physician Ratings

Disputes over the use and accuracy of performance ratings have arisen frequently in recent years, occasionally resulting in litigation. Some health care providers, for example, have brought defamation or libel claims against former patients who used the Internet to disseminate negative reviews of their care. Many recent legal developments related to quality reporting, however, have revolved around disputes between health care providers and health plans.

A. Disputes over Health Plan Rating Programs

Health plans have access to large volumes of data about their network physicians. From this data, they can extract cost and quality information for the purpose of creating "high-performance" networks. Physician tiering algorithms differ across plans. Historically, some plans have focused primarily on costs, partly because of the difficulty in assessing quality; others first apply a quality threshold and then assess efficiency. Quality measures can be based on practices


91. Ashish K. Jha et al., The Inverse Relationship Between Mortality Rates and Performance in the Hospital Quality Alliance Measures, 26 HEALTH AFF. 1104, 1104 (2007) ("The relationship between high HQA performance and lower risk-adjusted mortality is an important validation for this national hospital quality rating program.").


93. See, e.g., Draper et al., supra note 27, at 1 (describing creation of high-performance networks); cf. Theo Francis, Insurers' Lists on Doctors Under Fire, WALL ST. J., Aug. 21, 2007, at D3 (describing major insurers' efforts to create networks that rate physicians); Jason Roberson, Insurers' Ratings Often Aren't Accurate, Doctors Complain, DALLAS MORNING NEWS, Feb. 10, 2008 (describing ratings systems used by Texas insurers).

such as prescribing beta blockers for heart attack patients, performing blood sugar tests for diabetics, and screening for breast and cervical cancers.\(^{95}\)

Some employer-sponsored health insurance packages offer enrollees a broad choice of physicians, including both physicians in a high-performance network and physicians outside of it, but highlight the high-performance network designation in plan directories.\(^{96}\) Star ratings may guide enrollee choices of physicians, just as they sometimes guide selections of restaurants or movies.\(^{97}\) Some employers encourage the choice of more highly rated physicians through lower co-payments.\(^{98}\) Under other plans, lower-rated physicians may be excluded from provider networks altogether.\(^{99}\)

Rating and tiering programs often strain relationships between health plans and physicians.\(^{100}\) In 2005, for example, UnitedHealthcare operated a pilot program under which it used claims data to assess patient treatment costs and evaluate quality, as indicated by measures such as complication rates and adherence to medical guide-


\(^{96}\) Draper et al., supra note 27, at 3.

\(^{97}\) Id. See also Draper et al., supra note 27, at 3. One recent plan created three tiers based on physician rankings, with co-payments of $15, $25, or $35. Jeffrey Krasner, Insurers Expand Use of Doctor Rankings, Boston Globe, June 10, 2008, available at http://www.boston.com/business/healthcare/articles/2008/06/10/insurers_expand_use_of_doctor_rankings/.

\(^{98}\) House Hearing, supra note 94; Nakashima, supra note 95.

Physicians who met United's criteria received a "star" designation. Medical societies, physician groups, and health systems in St. Louis, Missouri objected to United's plan to implement the program in their area, arguing in part that risk adjustment for illness severity was inadequate, that the designations interfered with physician-patient relationships, that the program rewarded low cost rather than quality, and that the program considered costs over which physicians had no control. United subsequently modified the rating process and ultimately agreed to work with a local hospital system to develop quality and efficiency measures.

More recent efforts to implement rating programs have resulted in litigation. In 2006, Regence BlueShield used claims data to create a "Select Network" consisting of physicians who met certain quality and efficiency criteria. After the insurer sent letters to patients of excluded physicians, explaining that their physicians did not meet the requisite quality and efficiency standards and that they needed to select other physicians if they wanted continued coverage, the Washington State Medical Association (WSMA) demanded an apology. Regence later sent a letter to its members apologizing for any "[implication] that health care professional(s) who were not invited to participate in the Regence Select Network do not provide quality care," and delayed implementation of the program. Nevertheless, WSMA subsequently filed suit seeking damages and an injunction, alleging that Regence had used outdated data and a flawed methodology to create


103. See Judith VandeWater, Doctors Level Charges at Health Insurer, ST. LOUIS POST-DISPATCH, May 26, 2005, at A1 (describing physician objections); see also Kazel, supra note 100 (same).


106. Id.

the network.\textsuperscript{108} Causes of action included defamation, breach of contract, and violation of Washington's Consumer Protection Act.\textsuperscript{109}

Other lawsuits have followed. In 2007, for example, the Fairfield County Medical Association sued health plans that had created rating programs designating some physicians as elite.\textsuperscript{110} Similarly, in 2008, the Massachusetts Medical Society sued an entity that purchases health insurance for Massachusetts state employees.\textsuperscript{111} It alleged that the entity implemented a tiering program that had mis-tiered physicians, "falsely conveying the message to patients, the medical profession, and the community that the Mis-Tiered Physicians provide lower quality and less cost-effective care than physicians who are placed into a better tier . . . . Defendants have insufficiently reliable or valid methodology to measure the quality of care provided . . . ."\textsuperscript{112} The plaintiffs' causes of action included violation of procedural due process rights, defamation, tortious interference with advantageous relationships, violation of the Consumer Protection Act, and breach of contract.\textsuperscript{113}

At the core of all of these suits are physicians' concerns about the validity of measures used, the accuracy of underlying data, and the conflation of quality and cost criteria in creating tiers. The Regence suit resulted in a settlement agreement that reflects many of these con-


\textsuperscript{109.} See RINN, supra note 108, at 8.


\textsuperscript{112.} See First Amended Complaint, supra note 111, at 2.

\textsuperscript{113.} Id. One author has enumerated numerous potential causes of action in tiering cases: breach of contract; defamation or libel; state unfair trade practices or consumer protection violations; unfair insurance practices act violations; tortious interference with contractual relations; fraud; and conspiracy. RINN, supra note 108, at 4-7. Another group of authors adds the following: violation of statutory or common law due process requirements; violation of federal laws regulating health plans; violation of laws applicable to sponsored health plans; restraint of trade; and civil rights laws violations. SARA ROSENBAUM ET AL., AN ASSESSMENT OF LEGAL ISSUES RAISED IN "HIGH PERFORMING" HEALTH PLAN QUALITY AND EFFICIENCY TIERING ARRANGEMENTS: CAN THE PATIENT BE SAVED? (2007), available at http://www.rwjf.org/files/research/physician-tiering102007.pdf.
cerns. Regence agreed that for two years, before adopting a program involving performance measurement, it would disclose its measurement methodology and seek input from WSMA on topics such as the timeliness of data, risk-adjustment issues ("comparability of physician practices and patient populations"), and the appropriateness of applying measures to a given physician ("physician attribution issues"). Regence further agreed to post on its provider website its methodology, including an explanation of the data used. It also promised to permit physicians to appeal their scores, and if the physician disagreed with the outcome of the internal appeal, to permit binding review by an independent external reviewer. Disclosure of tiering methodologies and the availability of opportunities for appeal are important first steps toward ensuring the quality of quality ratings.

B. Redesigning Ratings: A Public Regulator's Role

Litigation between physicians and health plans has not been the only force shaping the nature of rating programs. In 2007, New York's Office of the Attorney General took a series of actions that will likely exert an enduring influence over their future development.

1. The New York Attorney General Gets Involved

In the summer of 2007, New York's Office of the Attorney General issued letters to several major insurers concerning their physician ranking programs. These letters outlined concerns about the quality rating programs, warned of the risk that they would confuse the public, and requested extensive information regarding program operations. The letters asked insurers to explain their rating methodolo-

---


115. Id.

116. Id.


118. See sources cited supra note 117.
gies, their marketing practices, and their plans to comply with New
York laws regarding health plans' evaluations of health care profession-
als, among other things.\textsuperscript{119}

The letters' language implied that the Attorney General viewed
existing rating programs as potentially highly problematic. A letter
concerning an Aetna program, for example, stated that the "program
carries a significant risk of causing consumer confusion, if not decept-
on."\textsuperscript{120} It emphasized the Attorney General's commitment to trans-
parency, and observed that "the goal of transparency is defeated . . . if
the information provided is in itself inaccurate or misleading, or based
on flawed data."\textsuperscript{121} Before requesting information about Aetna's pro-
gram, the letter noted that poorly designed ratings may "undermin[e]
[consumers'] ability to choose the best doctors," that "[c]onsumers
may be encouraged to choose doctors because they are cheap rather
than because they are good," and that the insurer's "profit motive may
affect the accuracy of its quality and cost-effectiveness rankings
because high-quality doctors may cost Aetna more money."\textsuperscript{122}

The concerns underlying the letters are unquestionably valid. At
the same time, though, the letters fail to fully acknowledge the environ-
ment in which insurer rating programs were developed. First, while it
is true that poor-quality ratings can hinder efforts to choose the best
doctors, there is little reason to believe that, but for these ratings,
patients would have done so. Patients have little systematic evidence
of quality to guide their decisions. Second, while it is possible to
design a program that would lead patients to choose physicians
"because they are cheap rather than because they are good," there is
not necessarily an inconsistency between being "cheap" and being
"good."\textsuperscript{123} It is likely that some physicians are "cheaper" but just as
good as other physicians, or even better. Evidence compiled by
researchers shows that wide regional variation in health care costs is
not necessarily associated with quality, and, indeed, that lower-cost
areas may offer superior care.\textsuperscript{124} Finally, it may be true that insurers


\textsuperscript{120}. Aetna Letter, supra note 117, at 1.

\textsuperscript{121}. Id. at 2.

\textsuperscript{122}. Id. at 3.

\textsuperscript{123}. Furthermore, in a health care system in which individuals are ultimately
responsible for funding their own care, a few patients might prefer to choose
physicians because they are "cheap," even if they are not quite as "good."

\textsuperscript{124}. See, e.g., Katherine Baicker & Amitabh Chandra, Medicare Spending, the
Physician Workforce, and Beneficiaries' Quality of Care, 23 Health Aff. w4-184, w4-184
(or employers that fund health insurance, or employees facing low wages due to high health care costs) have something to gain if they can reduce health care costs. Again, however, while some high-quality physicians may be high-cost, others may not be as costly. The ideal insurer rating program would identify physicians who offer high-quality care efficiently.

By reviewing concerns about rating programs at length, but only briefly referring to their potential benefits, the letters conveyed the impression that the programs were inevitably seriously flawed. Admittedly, flaws are a very real concern. As the Aetna letter points out, claims data may omit potentially relevant clinical information, databases might be too small to compile meaningful ratings, and attributing care to a single physician is problematic when multiple physicians are involved. But the current state of health care is also a very real concern. A well-designed rating program has the potential to both improve quality and reduce costs by guiding patients to higher-quality and/or more efficient providers, and by providing incentives to providers to change their practices for the better. The Attorney General’s letters, and the extensive data demands they contained, could have had the unfortunate effect of chilling the development and implementation of such programs.

2. The New York Settlement Agreements

Fortunately, the letters do not appear to have had this effect. Instead, they led to a series of settlements. The first, with CIGNA, came just a few months after the letters were issued. The settle-

---

(2004) (examining quality measures such as provision of beta blockers after heart attacks and blood sugar monitoring for diabetics and concluding that “states with higher Medicare spending have lower-quality care”); Elliott S. Fisher et al., The Implications of Regional Variations in Medicare Spending. Part 2: Health Outcomes and Satisfaction with Care, 138 ANNALS INTERNAL MED. 288, 288 (2003) (“Medicare enrollees in higher-spending regions receive more care than those in lower-spending regions but do not have better health outcomes or satisfaction with care.”). The Fisher et al. study found that regional differences in spending were in part related to physician visits. Specifically, it found that “differences in spending were explained almost entirely by greater frequency of physician visits, more frequent use of specialist consultations, more frequent tests and minor procedures, and greater use of the hospital and intensive care unit,” but it “found no evidence to suggest that the pattern of practice observed in higher-spending regions led to improved survival, slower decline in functional status, or improved satisfaction with care.” Id. at 292-93.


ments appear to reflect the input of a variety of stakeholders. In the CIGNA settlement press release, the Attorney General’s office explained that “[b]y working together with CIGNA, consumer advocates, doctors, and medical societies, we have been able to create a new template for ranking programs that the entire industry can follow.” A later press release mentioned that the “model was created in consultation with, and is supported by, the American Medical Association and the Medical Society of the State of New York, along with a host of consumer advocacy groups including the Consumers Union and the National Partnership for Women & Families.” The consumer groups appeared to be especially supportive of transparent, high-quality ratings. An AARP representative, for example, was quoted as saying that “[c]onsumers need valid and reliable information to choose a doctor that meets their needs. AARP strongly supports efforts to give consumers this type of information . . . ”

While the letters requesting information on rating programs focused almost exclusively on their pitfalls, the settlement agreements acknowledge that “wide variation in the quality and cost-efficiency of care delivered by health care providers and professionals is well-documented.” They explain that “meaningful efforts to measure and publicly report the comparative quality of physician practice are needed to help consumers make informed choices,” and that “experience has shown that measuring and publicly reporting physicians’ performance based on quality and cost-efficiency supports provider efforts to improve their performance.” The remainder of the agreements then lay out a series of mechanisms for addressing the potential pitfalls of tiering programs—mechanisms that permit the programs to continue, but will help to ensure their quality.


127. Id.
130. Id.
131. CIGNA Settlement, supra note 126, at 1.
132. Id.
133. Id. at 3, 4.
The settlements embrace a model code for "physician performance measurement, reporting and tiering programs" based on the "core principles" of "accuracy and transparency of information, oversight of the process, and fairness in comparison of physicians." The accuracy and transparency provisions include numerous requirements with respect to the nature of performance measurement. For example, the agreements require that measures of cost-efficiency be calculated and disclosed separately from measures of quality. While they may in addition be combined to create a single ranking, the weight attributed to each portion must be disclosed. In short, parties to the agreements will no longer be able to designate a "high performance" tier that mixes quality and cost considerations, leaving consumers to wonder about what exactly "high performance" might mean. The agreements further specify that insurers "shall not conduct rankings based solely on cost-efficiency, but shall consider quality dimensions."

The agreements also attempt to influence insurers’ rating approaches by stating that insurers "should seek to achieve the goals of safe, timely, effective, efficient, equitable and patient-centered care," goals that have been advocated by the Institute of Medicine, and that they "should seek to include patient experience as a measure of patient-centeredness." Further, the insurers must use quality measures based on nationally recognized guidelines and "[w]here available . . . shall use measures endorsed by the National Quality Forum ("NQF") or other entities whose work in the area of physician quality performance is generally accepted in the health care industry"; if NQF measures are unavailable, measures endorsed by AQA and accreditors

134. Id. at 1, 4.
135. E.g., id. at 4.
136. Id.
137. Id. at 7.
138. Id. at 4.
139. See, e.g., BD. ON HEALTH CARE SERVS., supra note 69, at 1, 118-19 (listing aims for quality improvement).
140. CIGNA Settlement, supra note 126, at 5.
should be used. Other agreement provisions focus on issues such as risk adjustment, sample size, and data currency and accuracy.

The agreements emphasize the importance of transparency, stressing that the methodology and data used must be fully disclosed. They require insurers to explain to consumers the rating system and data limitations, and the fact that there is a risk of error. The insurers must also disclose detailed information about methodologies to physicians, while at the same time establishing an appeals process for physicians to challenge their performance ratings.

Finally, the agreements require insurers to contract with a monitoring entity known as a "Ratings Examiner," which "shall be a nationally-recognized standard-setting organization" approved by the Office of the Attorney General. It must be "national in scope, independent, and an Internal Revenue Code § 501(c)(3) organization, and shall have existing standards and collection processes that would enable the transparency and accuracy terms of this Agreement to be satisfied." The Ratings Examiner's task is to review the insurer's rating programs with respect to all settlement agreement provisions.

3. The Aftermath: Rating Programs Move Forward

By February 2009, numerous insurers had entered agreements with the New York Attorney General, and many committed to follow

---


142. E.g., CIGNA Settlement, supra note 126, at 5-6 (risk adjustment and sample size); id. at 10 (data).

143. See, e.g., id. at 7-8.

144. See, e.g., id. at 9-10.

145. Id. at 10-11.

146. Id. at 11.

147. See, e.g., id. at 10-11.
the model code's provisions not just in their New York plans, but also nationwide. By adopting widespread reforms to their rating systems, the insurers may have increased the likelihood of avoiding actions by other attorneys general, as well as by state legislatures. In 2008, Colorado adopted the Physician Designation Disclosure Act, which imposes accuracy and transparency requirements for physician ratings similar to those in the New York settlements, but other states have been less willing to intervene in the ratings process.

Another factor promoting nationwide reforms of health plan rating programs is the Patient Charter for Physician Performance Measurement, Reporting and Tiering Programs (Patient Charter), which was released in 2008. The Patient Charter was created by the Consumer-Purchaser Disclosure Project, "a group of leading employer, consumer, and labor organizations working toward a common goal to ensure that all Americans have access to publicly reported health care performance information." Groups endorsing the Patient Charter include AARP, AFL-CIO, the Pacific Business Group on Health (a group with members that "represent approximately 3,000,000 employees, retirees and their families and nearly $10 billion in annual health care expenditures"), and the National Partnership for Women and Families, which was among the organizations mentioned as supporters of the New York settlements.


The press release announcing the Patient Charter explained that there is "wide variation in physician performance yet there is little information available to help consumers find and choose those who provide the best quality care" and that "[p]rograms that measure and report on the performance of physicians are integral to reforming health care to improve health outcomes for patients, create a more efficient health care system, and ultimately expand access to health care."\textsuperscript{154} It stated that "[w]here embraced, the Patient Charter will ensure both consumers and physicians will be able to understand, trust and contribute to how health plans rate doctors' performance."\textsuperscript{155}

The Patient Charter calls for measures that are meaningful to consumers, that include both quality and cost-efficiency information, and that reflect the six aims of the Institute of Medicine.\textsuperscript{156} The Charter states that "[m]easures and methodology should be transparent and valid," that limitations should be disclosed, that certain aspects of the measurement process should be assessed against national standards, that data should be aggregated wherever feasible, and that the effects of rating programs should be evaluated.\textsuperscript{157} It advocates the use of measures based on national standards, including measures endorsed by the National Quality Forum, and, if such measures are absent, those endorsed by AQA, the National Committee for Quality Assurance (NCQA), the Joint Commission, and federal agencies.\textsuperscript{158} It requires programs to allow physicians to request reviews of results they believe are inaccurate.\textsuperscript{159} It also suggests that health plans ask an "independent health care quality standard-setting organization to review" their tiering programs.\textsuperscript{160} In short, the key features of the Patient Charter closely resemble those of the New York settlement agreements. Health plans that have signed on to the Patient Charter include the large insurers Aetna, CIGNA, UnitedHealthcare, and Wellpoint, among others.\textsuperscript{161}

\begin{footnotes}
\footnotetext{154}{Press Release, Consumer-Purchaser Disclosure Project, \textit{supra} note 150, at 2.}
\footnotetext{155} {Id.}
\footnotetext{156} {CONSUMER-PURCHASER DISCLOSURE PROJECT, PATIENT CHARTER FOR PHYSICIAN PERFORMANCE MEASUREMENT, REPORTING AND TIERING PROGRAMS: ENSURING TRANSPARENCY, FAIRNESS AND INDEPENDENT REVIEW (2008), available at http://healthcaredisclosure.org/docs/files/PatientCharter.pdf.}
\footnotetext{157} {Id.}
\footnotetext{158} {Id.}
\footnotetext{159} {Id.}
\footnotetext{160} {Id.}
\footnotetext{161} {CONSUMER-PURCHASER DISCLOSURE PROJECT, STATUS OF HEALTH PLANS AND OTHER ORGANIZATIONS COMMITTED TO FULFILLING THE PATIENT CHARTER (2008), available at http://healthcaredisclosure.org/docs/files/PatientCharterStatus.pdf.}
\end{footnotes}
NCQA, a health plan accreditation organization that issues its own report cards on health plans, has been approved as an independent rating program reviewer both by the Consumer-Purchaser Disclosure Project and by the New York Attorney General. It has now posted online its initial reviews of several New York health plans.

C. A Few Observations

This brief description of the evolution of health plan physician rating programs offers interesting insights into not just the development of performance measurement mechanisms, but also the health care regulatory process. First, it highlights a shift in attitude in the quality reporting area. Many of the earliest efforts by health plans to create tiers based on cost and quality elicited strong public objections from physicians, along with calls to halt the programs. Resistance to rating programs was widespread. Today, objections to rating programs remain frequent, as evidenced by recent lawsuits, media coverage, and the New York letters. The Regence settlement, however, hinted at program features that would permit health plans and physicians to find common ground, such as physician participation in measure development, transparent methodologies, and opportunities for appeal. The New York settlement agreements and the Patient Charter pushed health plan quality ratings forward by emphasizing their potential benefits, while at the same time creating a framework to ensure high-caliber quality ratings, transparency for physicians and consumers, and independent oversight. In short, the focus seems to have shifted toward collaborative development of mechanisms that will correct rating program flaws.

Second, the chronology highlights the contributions of multiple stakeholders to the development of quality reporting programs. Newspaper reports on conflicts such as the one in St. Louis often focus on the insurers that create rating programs and the provider groups that resist them. Reports about litigation naturally focus on the parties involved, which, again, are typically the report card producers and the

165. See id. (providing links to reports on extent to which various health plans meet requirements of New York settlement agreements).
rated professionals or their representatives. The New York letters were created by the Office of the Attorney General and sent directly to health plans operating high performance networks. But as the chronology demonstrates, other stakeholders, especially employer groups and consumer groups, have played major roles in shaping rating programs and their oversight. Health plans often develop rating programs at the behest of their employer-clients, and employers have supported the creation and implementation of the Patient Charter. Similarly, consumer advocacy groups played a role in both the New York settlements and the Patient Charter. These groups have shown a strong commitment to the diffusion of rating programs and, like physicians, want to ensure that these programs are of high quality.

Third, while consumers, employers, and physicians all have exerted influence over the structure of health plan tiering programs, the involvement of a regulator appears to have been key in achieving rapid and widespread reforms. The letters issued by the Office of the Attorney General of New York focused the attention of major health plans on the problematic aspects of health plan rating programs. The settlement agreements offered a platform for multiple stakeholders to offer their input into the creation of a specific framework to address these problems. The Attorney General’s legal powers will help to ensure that the framework is actually implemented.

One final observation is that the settlement agreements and Patient Charter offer much promise as an approach to regulate rating programs, but also have limitations. In focusing on the goals of accuracy, transparency, oversight, and fairness, these documents establish a useful general framework that will help to ensure that consumers realize the full benefits of quality reporting. Many of the agreements’ requirements address areas of longstanding concern. The requirement to separately identify cost and quality-related criteria, for example, may help to alleviate physicians’ concerns about ratings, while making measures more useful for patients. The requirement of the appointment of an independent ratings examiner takes advantage of third party expertise in program oversight and helps to avoid many of the problems that can arise if regulations are too detailed.166

At the same time, some agreement provisions may prove to be overly prescriptive. For example, why must plans consider quality as well as cost in developing performance ratings? It is important that

166. Admittedly, third-party auditors and accreditation organizations are not without their own difficulties. To the extent these organizations exercise discretion and have objectives that differ from those of regulators (or consumers), their decisions will deviate from those that regulators (or consumers) would have made.
performance ratings be accurate and not misleading. It would not be appropriate to label a rating based solely on cost as a "performance rating," because the label may mislead customers into thinking that quality has been factored into the evaluation. And, of course, as this Essay argues, there are many good reasons to develop and publish quality ratings. In an environment in which treatment costs vary considerably and good quality measures are not always available, however, it seems that there should be room for enterprising health plans to develop entirely cost-based performance measures. To be sure, such programs may not be popular with consumers, who would worry about potential cost-quality tradeoffs, and even if they were popular, a health plan could skirt the prohibition simply by weighting performance measures heavily toward cost-based components. Nevertheless, detailed standards remain a concern. The basic problem is one that plagues any form of standard-setting: the more detailed the requirements and prohibitions the agreements contain, the less room there will be for benefit-enhancing innovation sparked by the competitive process. 167

Ultimately, however, the New York settlement agreements and the Patient Charter represent an important step forward for health plan rating programs and for health care quality report cards more generally. In the future, health plan enrollees will be significantly more likely to gain access to helpful and reliable performance measures.

III. ENHANCING GAINS FROM QUALITY REPORTING

Health care quality report cards have significant potential to increase the quality of American health care, but that potential has only begun to be realized. Report card developers, government enti-

167. Another concern along these lines is that the agreements refer to specific standard-setting organizations to guide measure choice. On the one hand, this level of specificity can help boost rating program quality in the short term; the NQF and AQA are both expert, well-respected organizations. This approach also helps to achieve standardization, increasing the likelihood that provider ratings will be consistent across insurers and trusted by consumers. In addition, the New York settlement agreements leave at least some room to choose alternative measures. See, e.g., CIGNA Settlement, supra note 126, at 5 ("Where available, CIGNA shall use measures endorsed by the National Quality Forum ("NQF") or other entities whose work in the area of physician quality performance is generally accepted in the healthcare industry."). On the other hand, there is a risk that the agreements' specificity (for example, the requirement that AQA and accreditors' measures be used only if NQF-endorsed measures are unavailable) might stifle the development of alternative approaches to quality measurement or standard-setting. See id.
ties, regulators, payers, providers, researchers, and consumers can all take steps to enhance the gains from quality reporting.

Report card developers can obviously have a dramatic impact on the usefulness of report cards as a quality improvement tool for consumers, physicians, and others. Whether report card creators are government agencies, health plans, non-profit organizations, commercial organizations, or others, they should work to make report cards accurate and transparent, relevant to their targeted users, and user-friendly. They should work to create new measures, particularly outcome-based measures, that are relevant for patients with specific diseases or in need of specific types of treatment. They should focus on developing measures of physician or physician group quality, given that physicians are often patients' points of entry into the health care system. They should consider not only what information consumers do value, such as patient experience information, but also the information that more fully informed consumers would likely value, such as technical clinical quality measures. They should explain why the measures they use are important. They should publicize the availability of their report cards and facilitate potential users' efforts to find them.

Governments have many potential roles to play in supporting the development of effective report cards.168 Government entities should continue to publish report cards. Existing federal report cards on hospitals and nursing homes are potentially useful starting points for consumers. State report cards for specific types of care, like those in New York and Pennsylvania, are potentially helpful for consumers and providers alike and provide useful models for other conditions and other states. Even if state or local government entities lack the resources to create their own report card data-collection programs, they can use existing data to present information in a useful way to their residents. Recently, the Los Angeles County Supervisors asked California state officials to require California nursing homes to display the star ratings available on Medicare's Nursing Home Compare website, as well as to include rating system information in admissions agreements.169 This is one way to raise public awareness of quality ratings and to encourage greater responsiveness of nursing homes to Medicare quality indicators.

168. For a fuller discussion of the role of government organizations in quality improvement, see Madison, supra note 7, at 1631-51.

Government entities can also promote health care quality ratings by strengthening the foundations of quality measurement systems. They can fund research on quality measurement and support measurement standardization efforts, where appropriate. They can promote the adoption of electronic medical records, which will greatly facilitate efforts to collect the detailed and comprehensive data necessary to create high-quality performance measures. Finally, they can search for ways to make data available to entities wishing to use it to create performance ratings. The fragmented nature of the health care system makes it difficult for any single entity to gather enough data to produce meaningful quality ratings. By pooling public and private payer data, more ratings become possible. While the federal government has emphasized the importance of transparency and advocated quality measurement initiatives, it has not always acted in full support of these goals. Citing concerns about physician privacy and a permanent injunction issued in the 1970s, HHS has resisted efforts by various entities to obtain Medicare data; most recently, it appealed (successfully) a federal district court ruling that had required it to turn over data pursuant to a Freedom of Information Act request. It is of course important to consider privacy concerns—particularly patients' privacy concerns—in constructing frameworks for sharing data, but overcoming such barriers is a key step toward improving health care quality reporting. If necessary, Congress should pass legislation permitting sharing of Medicare data for the purposes of developing rating programs.


171. See Consumers' Checkbook, Ctr. for Study of Servs. v. U.S. Dep't of Health & Human Servs., 554 F.3d 1046, 1049-51 (D.C. Cir. 2009) (noting that HHS argued that physician claims data was exempt from the Freedom of Information Act under a provision barring disclosure of certain files that would constitute a "clearly unwarranted invasion of personal privacy"); Robert Pear, Employers Push White House to Disclose Medicare Data, N.Y. TIMES, Apr. 11, 2006, at A1, available at http://www.nytimes.com/2006/04/11/washington/11medicare.html (reporting that Medicare had refused to share data with Business Roundtable, which represents many large companies). The Court of Appeals for the District of Columbia found that "physicians have a substantial privacy interest in the total payments they receive from Medicare for covered services." Consumers' Checkbook, 554 F.3d at 1051. The court also found that "the requested data does not serve any FOIA-related public interest in disclosure." Id. at 1056. It therefore concluded that the requested information was exempt from disclosure. Id.

172. Some employers have advocated for such legislation, and relevant bills were introduced in Congress in 2006 and 2007. See Draper et al., supra note 27, at 6 & n.5
As illustrated by Part II, regulators also have an important role to play in supporting the development of high-quality quality ratings. Most obviously, public regulators such as the New York Attorney General can step in to ensure that ratings do not mislead consumers. Private regulators, such as the NCQA, can assist in this task. Private entities such as large employers can take on similar regulatory functions. For example, they can insist that any health care quality ratings provided to their employees are based on solid methodologies and accurate data.

The New York settlement agreements and the Patient Charter offer a good starting point for those seeking to shape reporting programs. Another interesting proposal for regulation of health care quality ratings is to create a panel that would “define standards for measuring and reporting quality of care,” “describe the optimal training and certification needed to measure and report quality of care,” “design an auditing system,” and “develop a system to ensure that health care organizations are held accountable for the quality of care they report.” Regardless of the specific regulatory approach taken to encourage accuracy of quality rating, it is important that it permit continued experimentation in quality measurement and reporting.

Entities that fund the provision of health care, including public payers such as Medicare and Medicaid as well as private employers, can guide patients to higher-quality providers through the information they provide to beneficiaries. Benefit designs such as tiering can draw additional attention to quality issues. Payers can also potentially significantly increase the returns to quality reporting by tying payment to performance on quality measures. Pay-for-performance programs have become very popular, although evidence of their impact on quality is still limited. In theory, however, paying more for higher quality will provide an even greater incentive to improve along measured dimensions than that provided by public reporting alone.

Health care providers can increase the benefits of quality reporting in multiple ways.¹⁷⁵ First, providers can take advantage of performance measurements of all kinds to improve their own care. If a quality measure suggests that their care is below average, then they should identify explanations for the low rating; if the rating stems from deficiencies in their delivery of care, they can strive to correct them. Providers able to achieve above-average quality measures through innovative delivery systems can be emulated by other providers. Second, providers can collaborate to create performance measures appropriate for their own use in improving care or making referrals, as well as measures appropriate for patient use in selecting providers. Physician knowledge and experience are critical inputs into the development of high-caliber quality ratings, and yet many specialty societies have been slow to become involved in quality measurement activities.¹⁷⁶ By further expanding their involvement in quality reporting, providers can increase the likelihood that measures are both accurate and useful. Third, providers can help patients understand and use health care quality report cards by alerting patients to their existence and by explaining both their advantages and their limitations.

Researchers can improve health care quality report cards by continuing to develop health care quality measures. More treatment-specific and condition-specific quality measures are needed, as are more outcome measures of all types. Researchers can also take a closer look at attribution issues. Physicians are rightfully concerned about being subjected to performance measures over which they have little or no control. In some situations, however, it may make sense to develop quality measures that capture not what a specific provider actually controls, but instead what the provider could influence in an optimally structured delivery system.¹⁷⁷ A quality rating might be assigned, for example, to a provider in a good position to coordinate care, even if

¹⁷⁵. For a more detailed discussion of physicians' professional obligations in a world in which information about provider quality is available, see Kristin Madison & Mark Hall, Quality Regulation in the Information Age: Challenges for Medical Professionalism, in MEDICAL PROFESSIONALISM IN A NEW INFORMATION AGE, supra note 1.

¹⁷⁶. See Timothy G. Ferris et al., Physician Specialty Societies and the Development of Physician Performance Measures, 26 HEALTH AFF. 1712, 1712 (2007) (stating that physician specialty societies "have not yet played a major role in the development of measures" and exploring barriers to and facilitators of specialty society involvement).

¹⁷⁷. The Institute of Medicine has similarly proposed "that measurement of the health care delivery system should not be impeded by the impossibility of first identifying an accountable actor or the perception that responsibility for care is outside one's realm of control." Bd. on HEALTH CARE SERVS., supra note 69, at 98. It observes that "one valuable and intended effect of the integrated measurement system
that provider does not engage in coordination activities. Doing so would inform patients about an important dimension of care, while at the same time encouraging providers to assume more responsibility for the dimension of quality that is measured. More research is required, however, to determine how such measures should be structured.

Researchers can also advance efforts to improve report cards by studying their use by consumers, physicians, and others. With respect to consumers, for example, we need to better understand the formats and types of content that are most effective at communicating important information. We also need to continue to study the effects of report cards on health care quality, health care costs, and health care disparities. With the appearance of new report cards come more opportunities to conduct studies of report cards' impact.

Finally, consumers and patients can greatly enhance the benefits of quality report cards simply by using them. They should be aware that quality differs across providers along a variety of dimensions, and should consider what kinds of provider characteristics they value. They should search a variety of reputable sources for information about these characteristics, keeping in mind that measures may differ. They should also consider that existing quality measures are unlikely to fully reflect provider quality, so a quality report card should be only one of the sources they consider. But it should certainly be one.

**Conclusion**

Much work remains to be done in developing appropriate health care quality measures, ensuring their accuracy and transparency, and encouraging their widespread use. Report card developers, government entities, regulators, payers, providers, researchers, and consumers all could potentially contribute to efforts to improve the quality of report cards and expand their impact. Given report cards' considerable potential as a tool to improve health care quality, they have reason to do so.