Working Paper Series
Congressional Budget Office
Washington, D.C.

Issues and Challenges in Measuring and Improving the Quality of Health Care

Tamara Beth Hayford
Congressional Budget Office
Tamara.Hayford@cbo.gov

Jared Lane Maeda
Congressional Budget Office
Jared.Maeda@cbo.gov

Working Paper 2017-10

December 2017

The authors thank the following staff of CBO: Jessica Banthin, Jeffrey Kling, Lyle Nelson, and David Weaver for their helpful suggestions and guidance; Justin Lee for fact-checking; and Benjamin Plotinsky for editing. The authors would also like to thank Philip Ellis, formerly of CBO, for helpful suggestions and guidance, as well as Kyle Redfield, formerly of CBO, for assisting with the extensive literature review and with organizing the project. The authors also appreciate helpful comments and suggestions from Ledia Tabor of the Medicare Payment Advisory Commission.

Abstract

Various stakeholders have made significant efforts to measure and improve health care quality, spurred by landmark reports issued over a decade ago that highlighted serious deficiencies. Most payers now require providers of care to report on aspects of quality as a way to measure their performance and hold them accountable for it. The most common types of initiatives to measure and improve health care quality are public reporting programs and pay-for-performance programs. Under public reporting programs, providers' performance on quality measures is publicly disseminated to help consumers make informed choices about their care (which may also motivate providers to improve their quality). In pay-for-performance programs, providers' quality scores directly affect their payments. Both types of initiatives use various information and financial incentives to encourage providers to follow evidence-based guidelines and processes, improve patients' experiences when receiving care, and improve clinical outcomes. Despite the growing use of quality measures, progress has been slow, and many deficiencies in quality persist. This paper provides an overview of the current state of quality measurement, and it uses initiatives developed and implemented through the Medicare program to illustrate the key issues and challenges that arise in measuring and improving the quality of providers.

Keywords: health care quality, quality measurement, public reporting, pay for performance, value-based purchasing, providers, Medicare

JEL Classification: I18, H51

Contents

1. Introduction

Efforts to measure and improve the quality of health care have proliferated in recent years. Many payers now require providers to report on quality as a way to make them accountable for their performance.[1] Some states and other third parties have also developed programs to monitor and report on provider quality. Those initiatives use various quality measures and incentives to encourage providers to follow evidence-based guidelines, enhance patients' experiences, and improve clinical outcomes. This paper provides an overview of the current state of quality measurement and discusses the key issues and challenges in measuring provider quality and designing programs to improve it.[2]

Understanding the challenges in both measuring the quality of health care and developing programs to improve it has become increasingly important for the Congressional Budget Office, because many policymakers seek to reorient federal programs toward paying for the value rather than just the volume of health care services. For example, Medicare has already implemented numerous programs designed to improve the quality of care, including those that simply measure the quality of care delivered by particular providers and report that information to beneficiaries and those that modify the payments to providers on the basis of their measured quality. Moreover, Medicare intends to significantly expand the proportion of services that are provided through alternative payment models in which payments to providers depend partly on their quality or in which those payments depend on a combination of providers' performance on quality, resource use, clinical improvement activities, and the use of electronic medical records to report on measures.[3] Such efforts could have important effects on the quality of care received by Medicare beneficiaries and on Medicare spending under current law or future proposals—and could have broader effects on the U.S. health care system. However, those efforts could also have unintended consequences, such as encouraging providers to improve their ranking by avoiding sicker patients.[4] In this paper, we discuss the issues and trade-offs inherent in developing quality measures and implementing quality improvement initiatives, using programs developed by Medicare to illustrate those concepts.

---

[1] Christine Cassel and others, "Getting More Performance From Performance Measurement," *The New England Journal of Medicine*, vol. 371, no. 23, pp. 2145–2147 (December 2014), http://dx.doi.org/10.1056/NEJMp1408345.

[2] Although programs have been developed to measure and improve quality among providers and health plans, this paper focuses on quality programs aimed at providers.

[3] Centers for Medicare & Medicaid Services, *Quality Payment Program* (January 2017), https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/Quality-Payment-Program-Long-Version-Executive-Deck.pdf.

[4] Rachel M. Werner and David A. Asch, "The Unintended Consequences of Publicly Reporting Quality Information," *Journal of the American Medical Association*, vol. 293, no. 10 (March 2005), pp. 1239–1244, http://dx.doi.org/10.1001/jama.293.10.1239.

The State of Health Care Quality

Although serious deficiencies in the quality of health care delivered in the United States have probably existed for a long time, several reports issued over a decade ago brought shortcomings in quality into much sharper focus. Numerous studies found that there were large gaps between the health care that medical experts recommended and the care that people were actually receiving.[5] For example, many people who were treated in a doctor's office or hospitalized for particular conditions did not receive treatments that had been shown to be beneficial, and many did not receive recommended preventive services; other patients received services that had limited value. Studies from the 1990s also found that a significant proportion of medical procedures had been performed for inappropriate reasons and that some procedures or treatments were underutilized.[6] Moreover, a report by the Institute of Medicine (now known as the National Academy of Medicine) concluded that widespread errors in health care delivery caused a substantial number of patient injuries and deaths.[7] Studies also found that the quality of care varied greatly among geographic areas.[8] Those serious deficiencies and inconsistencies in the quality of care existed even though the United States was spending more per capita on health care than any other country.[9] Public and private payers have responded to the serious and widespread deficiencies in the quality of care by devoting substantial efforts to measuring and improving quality.

In some respects, the quality of care appears to have improved over the past decade, although serious problems remain. A recent report by the Agency for Healthcare Research and Quality (AHRQ) found that providers' performance on most of the quality indicators examined has improved over the past decade, whereas performance on nearly all of the others has remained

[5] For example, see Elizabeth A. McGlynn and others, "The Quality of Health Care Delivered to Adults in the United States," *The New England Journal of Medicine*, vol. 348, no. 26 (June 2003), pp. 2635–2645, http://dx.doi.org/10.1056/NEJMsa022615; and Mark A. Schuster, Elizabeth A. McGlynn, and Robert H. Brook, "How Good Is the Quality of Health Care in the United States?" *Milbank Quarterly*, vol. 76, no. 4 (December 1998), pp. 517–563, www.milbank.org/quarterly/articles/how-good-is-the-quality-of-health-care-in-the-united-states.

[6] Elizabeth A. McGlynn, "Assessing the Appropriateness of Care: How Much Is Too Much?" (RAND Corporation, 1998), www.rand.org/pubs/research_briefs/RB4522.html.

[7] Institute of Medicine, "To Err Is Human: Building a Safer Health System" (1999), https://iom.nationalacademies.org/Reports/1999/To-Err-is-Human-Building-A-Safer-Health-System.aspx.

[8] Elliott S. Fisher and John E. Wennberg, "Health Care Quality, Geographic Variations, and the Challenge of Supply-Sensitive Care," *Perspectives in Biology and Medicine*, vol. 46, no. 1 (Winter 2003), pp. 69–79, http://dx.doi.org/10.1353/pbm.2003.0004; and Stephen Jencks and others, "Quality of Medical Care Delivered to Medicare Beneficiaries: A Profile at State and National Levels," *Journal of the American Medical Association*, vol. 284, no. 13 (October 2000), pp. 1670–1676, http://dx.doi.org/10.1001/jama.284.13.1670.

[9] Gerard F. Anderson, Bianca K. Frogner, and Uwe E. Reinhardt, "Health Spending in OECD Countries in 2004: An Update," *Health Affairs*, vol. 26, no. 5 (September/October 2007), pp. 1481–1489, http://dx.doi.org/10.1377/hlthaff.26.5.1481.

about the same.[10] Specifically, nearly 60 percent of the quality measures examined showed improvement, about a third of quality measures showed no change, and 10 percent of quality measures showed worsening performance. The report concluded that although there has been significant progress in some areas, there are other areas that need more attention.[11] Drawing conclusions about the quality of care and how it has changed over time is challenging, however, because existing measures have limitations and because there are important dimensions of quality that are not captured by currently available measures.[12]

Definition of Quality in Health Care

Quality of care can be difficult to define, but a number of organizations and observers have proposed definitions. The Institute of Medicine (IOM) has defined quality as "the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge."[13] According to another definition, high-quality care is "providing patients with appropriate services in a technically competent manner, with good communication, shared decision making, and cultural sensitivity."[14]

In its report *Crossing the Quality Chasm*, the IOM recommended six aims for improving the health care system to address deficiencies in the quality of care and other limitations: health care should be *safe*, *effective*, *patient-centered*, *timely*, *efficient*, and *equitable*.[15] By the IOM's definition, safe care avoids injuries to patients; effective care is based on scientific evidence; patient-centered care is respectful of and responsive to patients' preferences, needs, and values; timely care is provided with minimal delays; efficient care avoids wasting resources; and equitable care does not vary because of personal characteristics such as sex, ethnicity, geographic location, and socioeconomic status. Not all of those goals focus on quality or have been used to develop quality measures for individual providers. In particular, although the

---

[10] Agency for Healthcare Research and Quality, *2015 National Healthcare Quality and Disparities Report and 5th Anniversary Update on the National Quality Strategy*, Pub. No. 16-0015 (April 2016), https://www.ahrq.gov/research/findings/nhqrdr/nhqdr15/index.html.

[11] Ibid., page 26.

[12] For example, see Robert M. Wachter, "Why Diagnostic Errors Don't Get Any Respect—and What Can Be Done About Them," *Health Affairs*, vol. 29, no. 9 (September 2010), pp. 1605–1610, http://dx.doi.org/10.1377/hlthaff.2009.0513; and Robert A. Berenson and Deborah R. Kaye, "Grading a Physician's Value—the Misapplication of Performance Measurement," *The New England Journal of Medicine*, vol. 369, no. 22 (November 2013), pp. 2079–2081, http://dx.doi.org/10.1056/NEJMp1312287.

[13] Kathleen N. Lohr, ed., *Medicare: A Strategy for Quality Assurance*, vol. 1 (National Academies Press, 1990), www.nap.edu/catalog.php?record_id=1547.

[14] Mark A. Schuster, Elizabeth A. McGlynn, and Robert H. Brook, "How Good Is the Quality of Health Care in the United States?" *Milbank Quarterly*, vol. 76, no. 4 (December 1998), pp. 517–563, www.milbank.org/quarterly/articles/how-good-is-the-quality-of-health-care-in-the-united-states.

[15] Institute of Medicine, *Crossing the Quality Chasm: A New Health System for the 21st Century* (2001), www.nap.edu/catalog/10027/crossing-the-quality-chasm-a-new-health-system-for-the.

concept of efficiency takes into account both the quality of care provided and the resources used to produce it, existing efficiency measures tend to evaluate spending rather than quality. Moreover, although adopting equity as a goal may help improve quality for the population as a whole, it is not typically used to measure the performance of specific providers for specific populations.

Although there is a lot of overlap in the definitions of quality, consumers typically place more emphasis on certain aspects of quality than some of the expert definitions would imply. For example, although consumers care about the competence and skill of their providers, they might have a difficult time evaluating the technical aspects of care on the basis of their personal interactions. Therefore, when evaluating quality, consumers might place greater emphasis on a provider's communication skills, the provider's convenience and accessibility, the cleanliness of the health care setting, and whether the treatment or procedure improved their symptoms or condition—all of which may combine to determine their overall experience of receiving care.[16]

Initiatives to Measure and Improve Quality

Evidence of widespread deficiencies in the quality of care has prompted both public and private payers to develop programs to measure and improve quality and hold providers accountable for the quality of their care. Such programs have proliferated over the past decade.[17] The most common types of programs are public reporting programs and pay-for-performance (P4P) programs. In public reporting programs, the payer collects information on selected measures of quality and publishes each provider's quality scores. Such programs provide information to help consumers choose high-quality providers and to motivate providers to improve quality. In contrast, P4P programs motivate providers by translating quality scores into financial rewards and penalties.

2. Quality Measures

Quality measurement is at the core of both public reporting and P4P programs. Selecting appropriate quality measures to evaluate providers is an essential component of those programs. Different quality measures are used to evaluate different aspects of the health care delivery system, but each has its own strengths and limitations. Therefore, initiatives use many different types of quality measures—and sometimes combine them into composite measures—to evaluate quality in the health care system.

---

[16] National Opinion Research Center, *Finding Quality Doctors: How Americans Evaluate Provider Quality in the United States* (July 2014), www.apnorc.org/projects/Pages/finding-quality-doctors-how-americans-evaluate-provider-quality-in-the-united-states.aspx; and Shoshanna Sofaer and others, "What Do Consumers Want to Know About the Quality of Care in Hospitals?" *Health Services Research*, vol. 40, no. 6 (December 2005), pp. 2018–2036, http://dx.doi.org/10.1111/j.1475-6773.2005.00473.x.

[17] Christine K. Cassel and others, "Getting More Performance From Performance Measurement," *The New England Journal of Medicine*, vol. 371, no. 23, pp. 2145–2147 (December 2014), http://dx.doi.org/10.1056/NEJMp1408345.

Several public and private organizations contribute to the development of quality measures. The Centers for Medicare & Medicaid Services (CMS) and AHRQ are the main public institutions that develop quality measures. The Joint Commission, a private nonprofit establishment that accredits hospitals and other health care organizations, is another major developer of quality measures. Provider organizations and other groups may also be involved in developing such measures.[18] The National Quality Forum (NQF), a nonprofit organization that is sponsored through public and private sources, aims to improve health care quality through measurement across the nation.[19] It does so by reviewing the quality measures submitted by other organizations and endorsing the measures that meet its criteria. It also reevaluates existing measures periodically, and it may retire measures.[20] Various programs use the measures that the NQF has endorsed in an effort to improve the comparability of health care quality.

Types of Measures

Avedis Donabedian developed an early framework for the measurement and evaluation of quality in health care. It included three types of measures: structural, process, and outcome.[21] Many of the measures that initiatives use to measure and improve quality fall into those three broad categories. Since the Donabedian study, developers have designed additional types of quality measures, such as patient experience measures, which largely focus on the interpersonal aspects of the provider-patient relationship.

Structural Measures. Structural measures evaluate the capacity or conditions of the physical settings in which providers deliver care as well as the capabilities of those providers. Examples of structural measures of quality for hospitals include the number of nursing care hours per patient-day and the adoption of systems to enable clinicians to prescribe medications electronically.[22] Those measures capture a provider's capacity to deliver care, but they do not describe how that care is delivered. Structural measures are less commonly used in quality measurement programs, although they are commonly used to establish criteria that providers

[18] American College of Surgeons, "National Quality Forum Endorses Two ACS NSQIP Measures" (press release, February 7, 2012), www.facs.org/media/press-releases/2012/nqf0212.

[19] National Quality Forum, "NQF's Mission and Vision" (accessed November 22, 2017), www.qualityforum.org/About_NQF/Mission_and_Vision.aspx, and "Funding" (accessed November 22, 2017), www.qualityforum.org/About_NQF/Funding.aspx.

[20] National Quality Forum, "Maintenance of NQF-Endorsed Performance Measures" (accessed November 22, 2017), www.qualityforum.org/Measuring_Performance/Endorsed_Performance_Measures_Maintenance.aspx, and "NQF Retires 22 Measures" (February 13, 2012), www.qualityforum.org/News_And_Resources/Press_Releases/2012/NQF_Retires_22_Measures.aspx.

[21] Avedis Donabedian, "Evaluating the Quality of Medical Care," *The Milbank Memorial Fund Quarterly*, vol. 44, no. 3, part 2 (July 1966), pp. 166–206, http://dx.doi.org/10.1111/j.1468-0009.2005.00397.x.

[22] National Quality Forum, "ABCs of Measurement" (accessed November 22, 2017), www.qualityforum.org/Measuring_Performance/ABCs_of_Measurement.aspx.

must meet in order to participate in Medicare and Medicaid. In that light, they could be considered necessary but not sufficient criteria for providing high-quality care.

Process Measures. Process measures reflect whether or not providers take specific actions that are consistent with professional standards of care and evidence-based guidelines for diagnosis, monitoring, and treatment. Examples include giving an antibiotic to surgical patients to help prevent infections, providing written educational material during stroke hospitalizations to patients or their caregivers about stroke care and prevention of recurrence, and documenting all of a patient's current medications and dosages in the medical record.[23] Process measures are the most common type of quality measure endorsed by NQF and also are frequently used in quality improvement initiatives.[24]

Process measures seek to measure the underuse, overuse, or misuse of care.[25] Underuse is the failure to deliver recommended health care services. Most process measures reflect whether recommended care was delivered and thus indicate the presence or absence of underuse. An example is the set of measures related to prescribing aspirin for heart attack patients, which were developed as process measures on the basis of research findings that doing so was associated with better outcomes.[26]

Overuse is the provision of services that are of low value—that is, likely to have little clinical benefit or have a greater likelihood of producing harm than benefits.[27] For example, researchers have highlighted the issue of antibiotic overuse—namely, prescribing antibiotics for viral illnesses—both because patients could experience unnecessary adverse drug reactions and because of the public health threat of antibiotic resistance.[28] Other services, such as computerized tomography (CT) scans for lower-back pain, are costly and can incidentally reveal abnormalities that would have otherwise gone undetected but whose treatment is not medically

---

[23] Ibid.

[24] National Quality Forum, "Quality Positioning System" (accessed July 13, 2015), www.qualityforum.org/QPS/QPSTool.aspx.

[25] Minal S. Kale and others, "Trends in the Overuse of Ambulatory Health Care Services in the United States," *JAMA Internal Medicine*, vol. 173, no. 2 (January 2013), pp. 142–148, http://dx.doi.org/10.1001/2013.jamainternmed.1022.

[26] R. M. Gunnar and others, "ACC/AHA Guidelines for the Early Management of Patients With Acute Myocardial Infarction," *Journal of the American College of Cardiology*, vol. 82, no. 2 (August 1990), pp. 664–707, www.ncbi.nlm.nih.gov/pubmed/2197021.

[27] Medicare Payment Advisory Commission, "Quality of Care in the Medicare Program," in *Health Care Spending and the Medicare Program* (June 2016), pp. 41–52, www.medpac.gov/docs/default-source/data-book/june-2016-data-book-health-care-spending-and-the-medicare-program.pdf?sfvrsn=0.

[28] Yuting Zhang, Michael A. Steinman, and Cameron M. Kaplan, "Geographic Variation in Outpatient Antibiotic Prescribing Among Older Adults," *Archives of Internal Medicine*, vol. 172, no. 19 (October 2012), pp. 1465–1471, http://dx.doi.org/10.1001/archinternmed.2012.3717.

necessary and could have harmful repercussions.[29] CT scans both introduce the risks associated with radiation exposure and often lead to surgery, which is generally no more successful at relieving pain than less invasive treatments, such as physical therapy.[30] NQF has endorsed various overuse measures, but many researchers and clinicians believe that overuse continues to account for a substantial portion of health care spending and advocate the development of additional measures to quantify it.[31] The American Board of Internal Medicine launched the "Choosing Wisely" campaign in 2012 in an effort to identify tests and procedures that were overused in a broad range of specialties and to encourage patients and clinicians to discuss those services and reduce their use when unnecessary.

Misuse is the delivery of the wrong type of care or the inappropriate delivery of care. Some examples are prescribing the wrong dosage of a medication and prescribing a medication to which the patient has a known allergy.[32] Although misuse is widely acknowledged as a problem in health care, the methods for identifying it are underdeveloped, and researchers have not reached a consensus on evaluating its occurrence.[33] Therefore, only a few of the process measures currently used focus on identifying misuse.

Outcome Measures. While process measures focus on the specific actions taken by the provider, outcome measures try to capture the effects of care delivery on a patient's health. Some outcome measures assess the health status of people with chronic conditions, such as blood glucose levels for diabetic patients and blood pressure levels for hypertensive patients. That type of measure, often called an "intermediate outcome," helps identify whether patients with chronic conditions are being managed well and are having their symptoms controlled effectively—without which they are much more likely to develop complications. Other outcome measures evaluate the frequency of adverse health events, such as mortality rates within 30 days of hospital admission and unplanned readmissions within 30 days of hospital discharge—both of which are often measured for patients with certain conditions or who have undergone certain procedures.

---

[29] National Quality Forum, *Measurement Framework: Evaluating Efficiency Across Patient-Focused Episodes of Care* (January 2010), www.qualityforum.org/Publications/2010/01/Measurement_Framework__Evaluating_Efficiency_Across_Patient-Focused_Episodes_of_Care.aspx.

[30] Choosing Wisely, "Imaging Tests for Lower-Back Pain" (April 2012), www.choosingwisely.org/patient-resources/imaging-tests-for-back-pain.

[31] Deborah Korenstein and others, "Overuse of Health Care Services in the United States: An Understudied Problem," *Archives of Internal Medicine*, vol. 172, no. 2 (January 2012), pp. 171–178, www.ncbi.nlm.nih.gov/pubmed/22271125.

[32] Ibid.

[33] Mathew Mercuri and Amiram Gafni, "Medical Practice Variations: What the Literature Tells Us (or Does Not) About What Are Warranted and Unwarranted Variations," *Journal of Evaluation in Clinical Practice*, vol. 17, no. 4 (April 2011), pp. 671–677, www.ncbi.nlm.nih.gov/pubmed/21501341.

Two other types of outcome measures have been developed more recently: patient safety measures and patient-reported outcomes. Patient safety measures capture the rate at which complications and adverse events occur following surgery or the delivery of other types of care. Examples include the rates of hospital-acquired infections, postoperative bleeding or bruising, and postoperative respiratory failure.[34] Patient-reported outcome measures are intended to capture improvements in health and functional status—such as gains in mobility after knee surgery or reductions in chest pain or other symptoms after cardiac procedures—that may be difficult to measure using claims data. Adopting such measures may also create incentives for providers to make care more patient-centered and to better engage patients in their care.[35] While quality initiatives frequently include patient safety measures in their measurement sets, they are just beginning to include patient-reported outcome measures.

Composite and Other Measures. There are a growing number of quality measures that do not fall neatly into the categories specified in the original Donabedian framework. Those mentioned here are composites of underlying existing measures or capture patient experiences.

Composite measures combine a number of traditional quality measures—weighted by an assessment of their relative importance—to create a single measure of quality.[36] For example, AHRQ's Patient Safety composite measure is composed of several individual patient safety measures, such as the rates of postoperative sepsis, accidental puncture or laceration, and postoperative respiratory failure, among others, which are weighted and combined into the overall measure.[37] Other composite measures, such as the CMS hospital star rating system, summarize several types of underlying quality measures in order to produce an overall quality rating for providers.

Patient experience measures, like patient-reported outcome measures, are a type of measure used to encourage providers to focus on patient-centeredness in care delivery, but they largely focus on the interpersonal aspects of the provider-patient relationship. That type of measure relates to patients' perceptions of how they were treated during care delivery, such as timeliness of appointments, providers' communication skills, and whether providers followed up on test

---

[34] National Quality Forum, *National Voluntary Consensus Standards for Public Reporting of Patient Safety Event Information* (February 2011), www.qualityforum.org/Publications/2011/02/National_Voluntary_Consensus_Standards_for_Public_Reporting_of_Patient_Safety_Event_Information.aspx.

[35] National Quality Forum, *Patient Reported Outcomes (PROs) in Performance Measurement* (January 2013), https://www.qualityforum.org/Publications/2012/12/Patient-Reported_Outcomes_in_Performance_Measurement.aspx.

[36] Lena Chen and others, "Composite Quality Measures for Common Inpatient Medical Conditions," *Medical Care*, vol. 51, no. 9 (September 2013), pp. 832–837, www.ncbi.nlm.nih.gov/pubmed/23942222.

[37] Agency for Healthcare Research and Quality, "AHRQ Quality Indicators: Composite Measures User Guide for the Patient Safety Indicators (PSI)" (September 2010), www.qualityindicators.ahrq.gov/Downloads/Modules/PSI/V42/Composite_User_Technical_Specification_PSI.pdf.

results.[38] Performing well on those measures would probably improve patients' experience and could also improve health outcomes through delivering care in a more timely fashion or ensuring that patients understand follow-up care instructions. For its Consumer Assessment of Healthcare Providers and Systems (CAHPS) program, AHRQ developed standardized survey tools that collect information on patients' experience with providers to help consumers select health plans and providers.[39] Several of the quality initiatives discussed in the following section incorporate CAHPS surveys in their quality measurement efforts.

Data Sources Used to Construct Quality Measures

The reliability of quality measures depends critically on the accuracy of the underlying data. Administrative data, medical records, registries, and surveys of patients are the main sources that reporting entities use to construct quality measures.[40] Each of those data sources has strengths and limitations for measuring quality.

Administrative data are derived from insurers' enrollment and claims files. The advantages of administrative data are that they are readily available and relatively inexpensive to obtain in electronic format for large populations. Administrative data are used to compute mortality and hospital readmission rates and certain other measures that require the diagnosis and procedure codes on claims. However, administrative data lack the clinical detail about patients' symptoms, medical history, complications, and test results that are required for many quality measures. Moreover, the accuracy and completeness of claims data vary by type of provider. In general, hospitals code patients' diagnoses more accurately and completely than physicians' offices do, because the payments from Medicare and many private insurers for inpatient care depend partly on patients' diagnoses, whereas payments to physicians typically depend on the services they provide but not on patients' diagnoses. Moreover, hospitals are audited and subject to financial penalties for incorrectly reported diagnoses.[41]

Administrative data are often used in conjunction with data from medical records because medical records include much more detailed clinical information. However, abstracting data from paper medical records requires expert staff and is much more expensive than obtaining information from administrative data. Moreover, medical records are less standardized than

---

[38]Agency for Healthcare Research and Quality, "Patient Experience Measures From the CAHPS Clinician & Group Surveys" (September 2014), www.ahrq.gov/sites/default/files/wysiwyg/cahps/surveys-guidance/cg/about/measures_cg.pdf.

[39] Agency for Healthcare Research and Quality, "About CAHPS" (October 2011), www.ahrq.gov/cahps/about-cahps/index.html.

[40] Robert A. Berenson, Peter J. Pronovost, and Harlan M. Krumholz, "Achieving the Potential of Health Care Performance Measures" (Robert Wood Johnson Foundation and Urban Institute, May 2013), www.rwjf.org/content/dam/farm/reports/reports/2013/rwjf406195.

[41] Agency for Healthcare Research and Quality, "Selecting Quality and Resource Use Measures: A Decision Guide for Community Quality Collaboratives. Part I: Introduction to Performance Data" (accessed November 22, 2017), www.ahrq.gov/professionals/quality-patient-safety/quality-resources/tools/perfmeasguide/perfmeaspt1a.html.

administrative data, and the quality of the data can vary among providers.[42] The adoption of electronic health records (EHRs) could increase the timeliness and lower the cost of accessing clinical information from medical records. However, the interoperability of different EHR systems is limited—making it difficult to collect data in a consistent way among organizations and settings—and EHRs have not been fully adopted by all providers.

Registries are another data source that can be used to evaluate the quality of providers. A patient registry is "an organized system that uses observational study methods to collect uniform data (clinical and other) to evaluate specified outcomes for a population defined by a particular disease, condition, or exposure, and that serves a predetermined scientific, clinical, or policy purpose(s)."[43] A limitation of registries is that they may not have standardized definitions or data elements; moreover, the data are not easily collected from different settings of care. Additionally, health information technology systems may not be interoperable with registries.

Patient surveys, such as CAHPS, are used to capture important information from patients' perspectives. The Medicare program requires hospitals and certain other providers to survey a sample of their patients during a specified period following discharge (using a standard set of questions and data collection methods) and levies payment penalties on those that fail to comply. The survey asks patients about their experiences with and ratings of their provider. A limitation of patient surveys is that they typically have low response rates, which raises questions about how well they represent the experiences of the entire target population. For example, the survey of patients that Medicare requires hospitals to conduct has a national average response rate of only 30 percent.[44] Such a low response rate raises concerns that the experiences of patients who responded to the survey might differ greatly from those of patients who did not respond. Moreover, differences in response rates could skew the comparisons of quality scores among providers.[45] Also, providers need to pay for the survey administration.

Strengths and Limitations of Quality Measurement

The advent of quality measurement has increased transparency in providers' delivery of health care. It has also allowed the identification of top performers in various areas—which, in

---

[42] Robert A. Berenson, Peter J. Pronovost, and Harlan M. Krumholz, "Achieving the Potential of Health Care Performance Measures" (Robert Wood Johnson Foundation and Urban Institute, May 2013), www.rwjf.org/content/dam/farm/reports/reports/2013/rwjf406195.

[43] Agency for Healthcare Research and Quality, *Registries for Evaluating Patient Outcomes: A User's Guide*, 3rd ed. (April 2014), https://effectivehealthcare.ahrq.gov/topics/registries-guide-3rd-edition/research.

[44] Centers for Medicare & Medicaid Services, "Summary of HCAHPS Survey Results: April 2014 to March 2015 Discharges" (December 2015), www.hcahpsonline.org/Files/December_2015_Summary_Analyses_Survey_Results.pdf.

[45] For example, see Zishan K. Siddiqui and others, "Comparison of Hospital Consumer Assessment of Healthcare Providers and Systems Patient Satisfaction Scores for Specialty Hospitals and General Medical Hospitals: Confounding Effect of Survey Response Rate," *Journal of Hospital Medicine*, vol. 9, no. 9 (September 2014), pp. 590–593.

conjunction with analysis of their methods, can help researchers and practitioners develop and refine best practices. For example, measurement and collaboration among clinicians in intensive care units have led to a marked reduction in central line–associated bloodstream infections. Measurement and the development of best practices in cardiovascular care have led to substantial reductions in hospitalization and mortality rates for acute myocardial infarction (heart attacks) and heart failure in the last several decades. A further benefit of quality measurement efforts has been an increased awareness of quality of care and thus further thought on the part of providers about improving their quality.[46]

While measurement has encouraged improvements in how health care is delivered, researchers have also identified weaknesses in the current state of quality measurement. One example is that existing measures tend to focus attention on conditions that are easily measured or clinical processes that are widely considered appropriate—a phenomenon sometimes called "looking under the streetlight." Although some quality-enhancing efforts (such as encouraging handwashing) may translate into broader quality improvements, other efforts may focus attention on measured aspects of care to the detriment of aspects that are not measured.

Similarly, the current set of quality measures cannot evaluate various complex clinical decisions and processes. For example, measures have not yet been developed to identify diagnostic errors, which are common and have significant potential to harm patients.[47] Existing measures also cannot assess physicians' ability to manage patients with multiple chronic conditions, nor can they capture other aspects of care, such as providers' technical proficiency or the accuracy of radiologists' interpretations of imaging studies.[48]

Also, some measures may have unintended consequences. For example, penalizing hospitals for cases in which patients fall down may lead some hospitals to encourage patients to stay in bed, which can reduce overall mobility and lead to blood clots, pressure ulcers, and other adverse outcomes.[49] In addition, stakeholders have expressed concerns that quality measures relating to

---

[46] Robert A. Berenson, Peter J. Pronovost, and Harlan M. Krumholz, "Achieving the Potential of Health Care Performance Measures" (Robert Wood Johnson Foundation and Urban Institute, May 2013), www.rwjf.org/content/dam/farm/reports/reports/2013/rwjf406195; and Haya R. Rubin, Peter J. Pronovost, and Gregory B. Diette, "The Advantages and Disadvantages of Process-Based Measures of Health Care Quality," *International Journal for Quality in Health Care*, vol. 13, no. 6 (December 2001), pp. 469–474, http://dx.doi.org/10.1093/intqhc/13.6.469.

[47] See Hardeep Singh and others, "Types and Origins of Diagnostic Errors in Primary Care Settings," *JAMA Internal Medicine*, vol. 173, no. 6 (March 2013), pp. 418–425, http://dx.doi.org/10.1001/jamainternmed.2013.2777.

[48] See Robert A. Berenson and Deborah R. Kaye, "Grading a Physician's Value—The Misapplication of Performance Measurement," *The New England Journal of Medicine*, vol. 369, no. 22 (November 28, 2013), pp. 2079–2081, http://dx.doi.org/10.1056/NEJMp1312287.

[49] Anna Gorman, "Pushing Elderly Patients to Keep Moving," *Modern Healthcare* (August 6, 2016), www.modernhealthcare.com/article/20160806/MAGAZINE/308069983.

pain management have encouraged the overprescription of painkillers.[50] For example, the Joint Commission established a pain standard in 2001 in response to advocacy efforts by pain specialists and other organizations in the late 1990s.[51] That standard may have contributed to the overprescription of opioids. A complicating factor, however, is that oxycodone was introduced around the same time and was heavily promoted.[52] Thus, a number of factors may have contributed to the problem.[53] CMS has recently responded to those concerns by removing questions relating to pain management from its survey of patients' experiences.[54]

Some researchers have also raised concerns about the gaming of quality measures that are self-reported. For example, hospitals might misreport data, such as hospital-acquired infection rates, or they might fail to report certain data to make their performance look better.[55] In addition, many quality measures focus only on what happens within a specific setting of care and not across the delivery system.

One way to improve quality measures would be to standardize their implementation. At present, different programs aiming to measure similar aspects of care have presented significantly different conclusions about the relative quality of different providers, sometimes even when the programs are using the same measures.[56] For instance, a hospital that ranks highly on the *U.S. News & World Report* "Best Hospitals" list may receive a poor mark on the CMS Hospital Compare rating and the associated Hospital Value-Based Purchasing Program rating. This type

---

[50] Virgil Dickson, "CMS Angers Hospitals With Plans for Site-Neutral Rates in Outpatient Payment Rule," *Modern Healthcare* (July 6, 2016), www.modernhealthcare.com/article/20160706/NEWS/160709964.

[51] David Baker, "The Joint Commission's Pain Standards: Origins and Evolution" (The Joint Commission, May 2017), www.jointcommission.org/assets/1/6/Pain_Std_History_Web_Version_05122017.pdf.

[52]Tatyana Lyapustina and G. Caleb Alexander, "The Prescription Opioid Addiction and Abuse Epidemic: How It Happened and What We Can Do About It," *The Pharmaceutical Journal* (June 11, 2015), www.pharmaceutical-journal.com/opinion/comment/the-prescription-opioid-addiction-and-abuse-epidemic-how-it-happened-and-what-we-can-do-about-it/20068579.article; and Art Van Zee, "The Promotion and Marketing of OxyContin: Commercial Triumph, Public Health Tragedy," *American Journal of Public Health*, vol. 99, no. 2 (February 2009), pp. 221–227, http://dx.doi.org/10.2105/AJPH.2007.131714.

[53] Jane Maxwell, "The Prescription Drug Epidemic in the United States: A Perfect Storm," *Drug and Alcohol Review*, vol. 30, no. 3 (May 2011), pp. 264–270, http://dx.doi.org/10.1111/j.1465-3362.2011.00291.x.

[54] Centers for Medicare & Medicaid Services, "CMS Finalizes Hospital Outpatient Prospective Payment System Changes to Better Support Hospitals and Physicians and Improve Patient Care" (press release, November 1, 2016), www.cms.gov/Newsroom/MediaReleaseDatabase/Press-releases/2016-Press-releases-items/2016-11-01.html.

[55] Department of Health and Human Services, Office of the Inspector General, "CMS Validated Hospital Inpatient Quality Reporting Program Data, But Should Use Additional Tools to Identify Gaming" (April 2017), https://oig.hhs.gov/oei/reports/oei-01-15-00320.pdf.

[56] Michael B. Rothberg and others, "Choosing the Best Hospital: The Limitations of Public Quality Reporting," *Health Affairs*, vol. 27, no. 6 (November/December 2008), pp. 1680–1687, http://dx.doi.org/10.1377/hlthaff.27.6.1680; and J. Matthew Austin and others, "National Hospital Ratings Systems Share Few Common Scores and May Generate Confusion Instead of Clarity," *Health Affairs*, vol. 34, no. 3 (March 2015), pp. 423–430, http://content.healthaffairs.org/content/34/3/423.abstract.

of inconsistency creates another layer of complexity for providers to muddle through as they attempt to respond to the incentives provided by the various quality initiatives.

Each of the main types of quality measures discussed above has different and counterbalancing strengths and limitations. Consequently, programs designed to measure and improve quality typically use a combination of measures that is intended to capture different aspects of care. The remainder of this section discusses the relative strengths and limitations of each type of measure.[57]

Structural Measures. The main advantages of structural measures are that they are fairly easy to report and compute and are straightforward to interpret. However, structural measures do not directly indicate whether recommended processes were followed or whether desired outcomes were achieved. There is also evidence of widespread variation in quality even among providers that meet certain minimum standards based on structural measures, which suggests that structural measures alone are not adequate for understanding and addressing that variation. Because of those limitations, structural measures are mostly used to determine whether providers meet certain minimum standards required for certification and accreditation.

Process Measures. Process measures have the advantage of identifying specific actions that clinicians can take to improve quality. They are typically based on professional standards of care and are therefore likely to be understood by physicians. Another advantage of process measures is that several of them are easy to calculate from administrative data, although some measures require administrators to abstract information from medical records, either to ascertain whether a procedure was performed or to determine whether a person should be included in the applicable pool of patients. Moreover, process measures do not usually need to be adjusted for patients' characteristics because success often does not depend on those characteristics—unlike outcome measures, which generally reflect both actions of providers and the extent and nature of the patients' illness. However, some process measures do depend on the actions of both providers and patients (such as the percentage of patients who actually receive certain preventive services), and some experts believe that those measures should be adjusted to account for differences in the socioeconomic status or other characteristics of patients.[58]

---

[57] The discussion of the advantages and disadvantages of different types of quality measures in this section draws on information from Robert A. Berenson, Peter J. Pronovost, and Harlan M. Krumholz, "Achieving the Potential of Health Care Performance Measures" (Robert Wood Johnson Foundation and Urban Institute, May 2013), www.rwjf.org/content/dam/farm/reports/reports/2013/rwjf406195; Haya R. Rubin, Peter J. Pronovost, and Gregory B. Diette, "The Advantages and Disadvantages of Process-Based Measures of Health Care Quality," *International Journal for Quality in Health Care*, vol. 13, no. 6 (December 2001), pp. 469–474, http://dx.doi.org/10.1093/intqhc/13.6.469; and Harlan M. Krumholz and others, "Measuring Performance for Treating Heart Attacks and Heart Failure: The Case for Outcomes Measurement," *Health Affairs*, vol. 26, no. 1 (January/February 2007), pp. 75–85, http://dx.doi.org/10.1377/hlthaff.26.1.75.

[58] National Quality Forum, *Risk Adjustment for Socioeconomic Status or Other Sociodemographic Factors* (August 2014),

An important disadvantage of process measures is that the evidence about whether higher scores lead to better outcomes is mixed. Some studies have found that when practitioners adhere to medically appropriate processes, mortality rates are lower.[59] However, others have noted a limited association between better processes and mortality.[60] One factor that may contribute to this issue is that process measures are generally computed for a subset of patients for whom a particular service is appropriate. For example, many of the process measures that CMS uses to evaluate the treatment of heart attack patients apply to fewer than half of those patients, while some apply only to about 10 percent—and it is challenging to identify whether performance on those measures is also associated with the quality of treatment for the unmeasured patients.[61]

Another disadvantage of process measures is that they might cause providers to focus resources on the particular processes being measured and divert resources away from other processes—a phenomenon sometimes called "teaching to the test"—which could result in no net change or possibly worse outcomes overall. In addition, process measures typically indicate whether or not a particular service was delivered, but not whether it was delivered effectively. For example, the measure of whether smoking-cessation counseling was performed does not take into account the duration or quality of the counseling. Many process measures are also "topped out"—meaning

---

www.qualityforum.org/Publications/2014/08/Risk_Adjustment_for_Socioeconomic_Status_or_Other_Sociodemographic_Factors.aspx.

[59] Rachel M. Werner and Eric T. Bradlow, "Public Reporting on Hospital Process Improvements Is Linked to Better Patient Outcomes," *Health Affairs*, vol. 29, no. 7 (July 2010), pp. 1319–1324, http://dx.doi.org/10.1377/hlthaff.2008.0770; and Rachel M. Werner and Eric T. Bradlow, "Relationship Between Medicare's Hospital Compare Performance Measures and Mortality Rates," *Journal of the American Medical Association*, vol. 296, no. 22 (December 2006), pp. 2694–2702, http://dx.doi.org/10.1001/jama.296.22.2694.

[60] Lauren Nicholas and others, "Hospital Process Compliance and Surgical Outcomes in Medicare Beneficiaries," *Archives of Surgery*, vol. 145, no. 10 (2010), pp. 999–1004, http://dx.doi.org/10.1001/archsurg.2010.191; Kyan Safavi and others, "Variation in Surgical Quality Measure Adherence Within Hospital Referral Regions: Do Publicly Reported Surgical Quality Measures Distinguish Among Hospitals That Patients Are Likely to Compare?" *Health Services Research*, vol. 49, no. 4 (August 2014), pp. 1108–1120, http://dx.doi.org/10.1111/1475-6773.12164; Ashish K. Jha and others, "The Inverse Relationship Between Mortality Rates and Performance in the Hospital Quality Alliance Measures," *Health Affairs*, vol. 26, no. 4 (July/August 2007), pp. 1104–1110, http://dx.doi.org/10.1377/hlthaff.26.4.1104; Andrew M. Ryan, Brahmajee K. Nallamothu, and Justin B. Dimick, "Medicare's Public Reporting Initiative on Hospital Quality Had Modest or No Impact on Mortality From Three Key Conditions," *Health Affairs*, vol. 31, no. 3 (March 2012), pp. 585–592, http://dx.doi.org/10.1377/hlthaff.2011.0719; and Andrew M. Ryan and others, "The Relationship Between Medicare's Process of Care Quality Measures and Mortality," *Inquiry*, vol. 46, no. 3 (Fall 2009), pp. 274–290, http://dx.doi.org/10.5034/inquiryjrnl_46.03.274.

[61] Robert A. Berenson, Peter J. Pronovost, and Harlan M. Krumholz, "Achieving the Potential of Health Care Performance Measures" (Robert Wood Johnson Foundation and Urban Institute, May 2013), www.rwjf.org/content/dam/farm/reports/reports/2013/rwjf406195.

that the average overall performance among providers is consistently high, making it difficult to meaningfully distinguish between low- and high-quality providers.[62]

Outcome Measures. The primary advantage of outcome measures is that they directly measure what patients care about—namely, whether their interaction with the health care system improved their health or led to an adverse event. Despite the appeal of measuring outcomes, the widespread adoption of such measures has been hampered by some important challenges.

A key challenge is how to adjust measures for underlying differences in health status—otherwise known as "risk adjustment." Unlike most process measures, most outcomes depend on the patient's health status as well as the provider's quality. For example, risk of readmissions or mortality following cardiac surgery will differ between patients who have other chronic conditions and those who do not. Current methods that adjust outcome measures for illness severity, comorbidities, and other characteristics of patients cannot perfectly predict relative risk across patients, which has two potential consequences. First, imperfect risk adjustment could make it difficult to differentiate between low- and high-quality providers. Second, if providers are better able to identify risk among patients than the risk-adjustment mechanism is, they might try to avoid certain high-risk patients.

Moreover, there is debate about whether certain measures should be adjusted for differences in patients' socioeconomic status. Proponents of adjusting for socioeconomic status argue that some outcomes depend in part on socioeconomic factors that might affect a patient's ability to follow a provider's treatment plan—such as an inability to afford a prescription medication or lack of transportation to a follow-up appointment—and that providers should not be penalized for worse outcomes that result from those factors. Opponents of adjusting for socioeconomic status are concerned that it would mask disparities in outcomes and amount to an acceptance of worse outcomes for patients of lower socioeconomic status.[63] In addition, measures of socioeconomic status such as income are not typically available in administrative data. One way to include socioeconomic status in a risk-adjustment system would be to use a geographic identifier, such as ZIP codes or counties in the administrative data, that could be linked with another data source to compute community-level measures of socioeconomic status. However, five-digit ZIP codes or counties often represent too large an area to yield reliable measures of socioeconomic status for individuals.[64]

---

[62] Charles N. Kahn III and others, "Assessing Medicare's Hospital Pay-for-Performance Programs and Whether They Are Achieving Their Goals," *Health Affairs*, vol. 34, no. 8 (August 2015), pp. 1281–1288, http://dx.doi.org/10.1377/hlthaff.2015.0158.

[63] For a discussion of these issues as it relates to risk adjustment for the hospital readmission rate, see Steven H. Sheingold and others, "Understanding Medicare Hospital Readmission Rates and Differing Penalties Between Safety-Net and Other Hospitals," *Health Affairs*, vol. 35, no. 1 (January 2016), pp. 124–131, http://dx.doi.org/10.1377/hlthaff.2015.0534.

[64] Susannah Bernheim and others, "Accounting for Patients' Socioeconomic Status Does Not Change Hospital Readmission Rates," *Health Affairs*, vol. 35, no. 8 (August 2016), pp. 1461–1470,

Even with effective risk adjustment, assigning appropriate outcome measures of quality to individual providers can be challenging. A particular challenge in measuring outcomes is that the number of patients who are included in the calculation might be too small to yield reliable results for many providers. This is a distinct concern for adverse outcomes that are rare, such as mortality rates for patients with certain conditions and postsurgical infection rates. An additional challenge is that many outcomes of interest—such as survival following cancer treatment—should be measured over a period of years. But such measures would reflect the care delivered by the provider years earlier and would not capture any changes in care processes that the provider might have implemented subsequently.[65] Consequently, such long-term outcomes might be of limited relevance for helping patients choose providers. Still another challenge in measuring outcomes is that some of them are influenced by the actions of many different providers, and in those cases it is often unclear which provider should be held primarily responsible. For example, many outcomes of the patients in a primary care group practice will be affected by the care they receive from a wide range of other providers.

In addition, certain outcome measures are subject to surveillance bias—the phenomenon in which providers that look more intensively for a particular outcome are more likely to find it. One commonly cited example is postoperative venous thromboembolism (blood clots), which requires imaging to diagnose. The same phenomenon may apply to adverse outcomes that are easier to identify, such as pressure ulcers (bedsores).[66] If providers vary in the methods and intensity with which they monitor such adverse outcomes, the reported differences in such measures will not provide a reliable indicator of quality differences; indeed, providers that are more fastidious about identifying adverse developments could be at a disadvantage.

Composite and Other Measures. Composite measures have the advantage of reducing consumers' cognitive overload by summarizing the results of multiple quality measures in a single rating. As a result, composite measures may be easier for consumers to understand and easier to display in reports. In addition, researchers have found composite measures to be more reliable than underlying individual measures in some narrowly defined areas, particularly for

---

http://dx.doi.org/10.1377/hlthaff.2015.0394; National Academy of Medicine, "Accounting for Social Risk Factors in Medicare Payment" (2016), www.nap.edu/catalog/21858/accounting-for-social-risk-factors-in-medicare-payment-identifying-social; National Quality Forum, *Evaluation of NQF's Trial Period for Risk Adjustment for Social Risk Factors* (June 2017), www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=85396; and Department of Health and Human Services, Assistant Secretary for Planning and Evaluation, *Social Risk Factors and Performance Under Medicare's Value-Based Purchasing Programs* (December 2016), https://aspe.hhs.gov/system/files/pdf/253971/ASPESESRTCfull.pdf.

[65] See Robert H. Brook, Elizabeth A. McGlynn, and Paul G. Shekelle, "Defining and Measuring Quality of Care: A Perspective from U.S. Researchers," *International Journal for Quality in Health Care*, vol. 12, no. 4 (August 2000), pp. 281–295, http://dx.doi.org/10.1093/intqhc/12.4.281.

[66] Karl Y. Bilimoria and others, "Evaluation of Surveillance Bias and the Validity of the Venous Thromboembolism Quality Measure," *Journal of the American Medical Association*, vol. 310, no. 14 (October 2013), pp. 1482–1489, http://jama.jamanetwork.com/article.aspx?articleid=1748150.

providers with a smaller sample of patients.[67] One study also found that an improvement in composite measures based on measures collected by CMS was causally associated with lower readmissions and mortality.[68] However, composite measures may mask variation among departments or types of measures (such as process or outcome measures). They may reflect performance on only a subset of conditions that a hospital treats.

Furthermore, it is challenging to construct an appropriate weighting system for the underlying measures.[69] That challenge underlies many of the criticisms of the composite measure that CMS constructed for hospitals—the overall hospital star rating system in Hospital Compare and the total performance score in the Hospital Value-Based Purchasing Program. CMS calculates a hospital's overall star rating as a weighted average of seven underlying measures and places greater weight on outcome measures and patient experience measures than on clinical process measures. However, smaller hospitals tend to be less likely to meet the sample size requirements for reporting on outcome measures. In those cases, more weight is placed on how those hospitals perform on clinical process measures, often leading to a higher star rating.[70]

In CMS's Hospital Value-Based Purchasing Program, hospitals with sufficient data for at least two out of four domains received a total performance score in fiscal years 2015 and 2016.[71] Hospitals with missing domain scores were reweighted proportionately on the basis of their scored domains. Under this approach, some hospitals with high efficiency scores but relatively low quality scores, especially some small rural and small urban hospitals, received bonuses. The efficiency score also had a greater effect for hospitals that were missing one or more quality scores. This method of scoring has raised concerns from some analysts because it contradicts

---

[67] Rebecca Lipner and others, "A Three-Part Model for Measuring Diabetes Care in Physician Practice," *Academic Medicine*, vol. 82, no. 10 (October 2007), pp. 48–52, www.ncbi.nlm.nih.gov/pubmed/17895690.

[68] Joseph J. Doyle, Jr., John A. Graces, and Jonathan Gruber, *Evaluating Measures of Hospital Quality*, Working Paper 23166 (National Bureau of Economic Research, February 2017), www.nber.org/papers/w23166.

[69] See Agency for Healthcare Research and Quality, "Combining Measures Into Composites or Summary Scores" (accessed February 2, 2017), www.ahrq.gov/professionals/quality-patient-safety/talkingquality/create/scores/combinemeasures.html; Francis J. Crosson, Medicare Payment Advisory Commission, letter to the Honorable Andrew Slavitt, Acting Administrator, Centers for Medicare & Medicaid Services (September 22, 2016); Karl Y. Bilimoria and Cynthia Barnard, "The New CMS Hospital Quality Star Ratings: The Stars Are Not Aligned," *Journal of the American Medical Association*, vol. 316, no. 17 (November 2016), pp. 1761–1762, http://dx.doi.org/10.1001/jama.2016.13679; and Susan Xu and Atul Grover, "CMS' Hospital Quality Star Ratings Fail to Pass the Common Sense Test," *Health Affairs Blog* (November 14, 2016), http://healthaffairs.org/blog/2016/11/14/cms-hospital-quality-star-ratings-fail-to-pass-the-common-sense-test/.

[70] Susan Xu and Atul Grover, "CMS' Hospital Quality Star Ratings Fail To Pass the Common Sense Test," *Health Affairs Blog* (November 14, 2016), http://healthaffairs.org/blog/2016/11/14/cms-hospital-quality-star-ratings-fail-to-pass-the-common-sense-test/.

[71] In fiscal years 2017 and 2018, CMS revised its policy to allow hospitals with sufficient data in at least three out of four domains to receive a total performance score. See Centers for Medicare & Medicaid Services, "Hospital Value-Based Purchasing" (September 2015), https://www.cms.gov/Outreach-and-Education/Medicare-Learning-Network-MLN/MLNProducts/downloads/Hospital_VBPurchasing_Fact_Sheet_ICN907664.pdf.

CMS's goal of rewarding hospitals that provide high quality at a lower cost.[72] These examples demonstrate the importance of carefully considering weighting systems for composite measures, particularly those that span multiple measure types.

Patient experience measures have the advantage of providing a better understanding of how treatment and care delivery affect the patient. In addition, patients' experiences capture a separate dimension of quality that does not simply reflect clinical outcomes, and they may help identify whether the care being delivered meets a patient's expectations. However, there are some challenges with accurately measuring patients' experiences. For example, although there are a growing number of patient experience measures being used, there is no consensus on a standardized definition, because patient experience is an ambiguous concept. As a result, some researchers have questioned whether existing patient experience measures are measuring the appropriate construct and whether they actually reflect care quality.[73] On the other hand, one study found that Yelp reviews covered the majority of domains in the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) as well as 12 domains that were not in HCAHPS.[74]

3. Initiatives to Measure and Improve Health Care Quality

There are two broad categories of programs to measure and improve health care quality. Public reporting programs publicly report providers' performance, which can encourage patients to use higher-quality providers and spur providers to improve their quality. Pay-for-performance programs incentivize providers' performance directly through financial rewards or penalties that are based on their quality scores.

Public Reporting Programs

Public reporting programs collect data on providers' performance for various types of quality measures and then report that information so that it is made broadly available to the public. Such programs are designed to serve two key functions: giving patients information to help them make informed choices about where to receive their care, and allowing providers to identify areas for improvement and motivating them to improve.[75] Providers' performance on quality measures, however, does not necessarily have to be reported publicly. In some cases, health care

[72] Government Accountability Office, "Hospital Value-Based Purchasing: CMS Should Take Steps to Ensure Lower Quality Hospitals Do Not Qualify for Bonuses" (June 2017), www.gao.gov/assets/690/685586.pdf.

[73] Sherri L. LaVela and Andrew S. Gallan, "Evaluation and Measurement of Patient Experience," *Patient Experience Journal*, vol. 1, no. 1 (April 2014), pp. 28–36, http://pxjournal.org/journal/vol1/iss1/5/.

[74] Benjamin L. Ranard and others, "Yelp Reviews of Hospital Care Can Supplement and Inform Traditional Surveys of the Patient Experience of Care," *Health Affairs*, vol. 35, no. 4 (April 2016), pp. 697–705, http://dx.doi.org/10.1377/hlthaff.2015.1030.

[75] Robert Wood Johnson Foundation, *Reform in Action: Can Publicly Reporting the Performance of Health Care Providers Spur Quality Improvement?* (August 2012), www.rwjf.org/content/dam/farm/reports/issue_briefs/2012/rwjf400299.

organizations report confidential feedback to providers for internal quality improvement purposes. That type of program supports the improvement goals of providers and health care organizations but does not support consumers in their decision-making process.[76]

The Medicare program reports information on the quality of care provided by hospitals, nursing homes, home health agencies, and dialysis facilities on its websites (see Table 1). The agency is in the process of expanding information reported on physicians.

Public reporting has the potential to improve health care quality by changing the behavior of consumers, providers, or both.[77] Consumers could respond to quality reports by choosing providers with higher quality ratings over those with lower ratings. Consumers may also shift to higher-quality providers if physicians change their referral patterns or private insurers change their selection of physicians for their networks. Both mechanisms would tend to increase the market share of high-quality providers and therefore could increase the average quality of care, even if individual providers do not improve their own quality.[78] The extent to which that occurs, however, depends on whether higher-quality providers have the capacity to treat more patients; if not, then such changes could merely rearrange which patients are treated by which doctors without having a substantial effect on care quality overall.

For their part, providers might work to improve their quality for two reasons. First, providers might worry about losing market share to others with higher quality scores. Second, providers might have an intrinsic motivation to deliver high-quality care and to maintain or improve their reputation.[79]

Some of the mechanisms through which public reporting programs could improve the quality of care depend on consumers' using quality scores to select providers. Consumers, however, have had limited engagement with public reports. Barriers include their lack of awareness, trust, and understanding of report cards. According to surveys, consumers are more likely to select a provider on the basis of their own past experiences, providers' reputations, or recommendations

---

[76] Agency for Healthcare Research and Quality, *Confidential Physician Feedback Reports: Designing for Optimal Impact on Performance* (March 2016), https://www.ahrq.gov/sites/default/files/publications/files/confidreportguide_0.pdf.

[77] Damien Contandriopoulos, François Champagne, and Jean-Louis Denis, "The Multiple Causal Pathways Between Performance Measures' Use and Effects," *Medical Care Research and Review*, vol. 71, no. 1 (February 2014), pp. 3–20, http://dx.doi.org/10.1177/1077558713496320.

[78] Judith H. Hibbard and others, "The Impact of a CAHPS Report on Employee Knowledge, Beliefs, and Decisions," *Medical Care Research and Review*, vol. 59, no. 1 (March 2002), pp. 104–116, http://dx.doi.org/10.1177/107755870205900106; and Rachel M. Werner and David A. Asch, "The Unintended Consequences of Publicly Reporting Quality Information," *Journal of the American Medical Association*, vol. 293, no. 10 (March 2005), pp. 1239–1244, http://dx.doi.org/10.1001/jama.293.10.1239.

[79] Donald M. Berwick, Brent James, and Molly Joel Coye, "Connections Between Quality Measurement and Improvement," *Medical Care*, vol. 41, no. 1 (January 2003), pp. I30–I38, http://journals.lww.com/lww-medicalcare/Abstract/2003/01001/Connections_Between_Quality_Measurement_and.4.aspx.

from family and friends.[80] Among consumers who are aware of public reports and would feel comfortable using them, the reports may not provide information that is useful or timely.[81] There is also limited evidence that public reporting programs have affected physician-referral patterns.[82]

In contrast, studies of providers' responses to public reporting have found that they engage in quality improvement activities when performance data are made publicly available—which suggests that their intrinsic motivations are the strongest factor. Both organizations and individual clinicians have been found to respond to public reports by making positive changes to their behavior. For example, some studies have shown that providers have responded to quality reports by increasing their quality improvement activities, while other studies suggest that surgeons with worse outcomes were more likely to stop practicing. Thus, public reporting has been found to be more strongly associated with changes in providers' behavior than with the selection of providers by patients.[83]

Pay-for-Performance Programs

In pay-for-performance programs, the payments that providers receive depend partly on their quality scores. P4P programs are intended to shift payments from a volume-based system to a value-based system. There are two broad categories of P4P programs: those that give providers financial incentives to meet quality benchmarks (that is, quality is an objective); and those that give providers financial incentives to reduce the cost of care but also require providers to meet specified quality targets in order to receive some or all of those rewards (that is, quality is a constraint).

P4P programs are intended to improve the quality of care primarily by affecting the behavior of providers. However, many P4P programs also include a public reporting component or operate in conjunction with a public reporting program, which might also affect quality by changing the behavior of consumers as discussed above. Payers could also affect the overall quality of care by using information from those reporting programs to modify cost-sharing requirements (for

[80] National Opinion Research Center, *Finding Quality Doctors: How Americans Evaluate Provider Quality in the United States* (July 2014), www.apnorc.org/PDFs/Finding%20Quality%20Doctors/Finding%20Quality%20Doctors%20Research%20Highlights.pdf.

[81] Jonathan T. Kolstad and Michael E. Chernew, "Quality and Consumer Decision Making in the Market for Health Insurance and Health Care Services," *Medical Care Research and Review*, vol. 66, no. 1 (February 2009), pp. 28S–52S, http://dx.doi.org/10.1177/1077558708325887.

[82] Ibid.

[83] Annette M. Totten and others, *Public Reporting as a Quality Improvement Strategy. Closing the Quality Gap: Revisiting the State of Science*, Evidence Report/Technology Assessment 208 (Agency for Healthcare Research and Quality, July 2012), www.ncbi.nlm.nih.gov/books/NBK99879/pdf/Bookshelf_NBK99879.pdf.

instance, through a tiered provider network), thus giving patients an incentive to select higher-quality and lower-cost providers.

The financial incentives in P4P programs that have quality as an objective can include penalties for providers that fail to meet quality targets, bonuses for providers that meet or exceed such targets, or a combination of both. Medicare has such P4P programs in place for hospitals, nursing homes, home health agencies, and renal dialysis facilities (see Table 2). Moreover, the recently enacted Medicare Access and CHIP Reauthorization Act (MACRA) modified Medicare's physician payment system to establish a new P4P program for physicians, called the Quality Payment Program, and it will replace existing programs, which we discuss below.[84]

The general design of P4P programs with quality as a constraint is that providers are allowed to share in the savings they generate if they hold health care spending for a defined set of patients below a certain target—but only if they meet or exceed certain quality targets. (In some cases, providers are also placed at risk for financial losses if their spending exceeds the target, whether or not they improve quality.) Insurers impose such a constraint to ensure that providers do not achieve savings in ways that reduce quality. A prominent example of such a program is the accountable care organization (ACO), in which doctors, hospitals, and other providers have an incentive to work together to coordinate care and reduce total spending for the patients attributed or assigned to them. Medicare is experimenting with ACOs in an effort to contain costs and requires those organizations to demonstrate that they met a quality performance standard in order to share in the savings they may have generated.

Issues and Challenges in Designing Programs to Measure and Improve Quality

Programs seeking to measure and improve the quality of health care can be designed in various ways. For example, program designers need to decide whether to include a broad or narrow set of measures and measure types as well as whether and how to adjust for differences in patients' health risk. Designers of public reporting programs need to decide how to present information to consumers, and designers of P4P programs need to decide on a structure for their financial incentives. More broadly, program designers also need to decide whether to allow providers to voluntarily participate in quality initiatives. This section discusses those key choices as well as the strengths and limitations of the various options.

Measure Selection. An important step in designing quality initiatives is determining which quality measures to include. Because each type of measure has strengths and limitations, many quality initiatives include a broad range of measures in order to capture different aspects of care. For example, as of July 2016, Medicare's public reporting program for hospitals presented 100 quality measures on the Medicare.gov website, including 8 structural measures, 48 process

[84] The Quality Payment Program creates two different tracks for physicians to participate in: the Merit-Based Incentive Payment System (MIPS) and alternative payment models (APMs). The MIPS replaces the Physician Quality Reporting System (PQRS), the Medicare EHR Incentive Program (or "meaningful use"), and the value-based modifier.

measures, 31 outcome measures, 12 patient experience measures (including a composite of the individual measures), and a new overall composite star rating based on 64 of the individual measures.[85] Medicare's Hospital Value-Based Purchasing Program applies bonuses or penalties to hospitals' payments depending on their scores on a subset of those measures. Both programs also include a measure of average spending per beneficiary among patients with selected conditions—reflecting an effort to measure value and not just quality.[86]

Including a broad range of measures may produce a more complete depiction of providers' quality, because different measures focus on different conditions or components of health care. The inclusion of a broad range of measures might also limit the ability of providers to "teach to the test" and focus only on measured quality rather than overall quality. Even so, a recent study found that hospitals improved quality more in areas that were more likely to increase their bonus payment from Medicare's Hospital Value-Based Purchasing Program, suggesting that hospitals may be able to track performance on a broad variety of measures in order to maximize revenues.[87]

Other programs use a narrower set of measures, reflecting their narrower focus on only certain dimensions of quality. For example, Medicare's readmission reduction program penalizes hospitals whose unplanned-readmission rates for selected conditions are relatively high, and Medicare has another pair of programs that penalize hospitals whose patients are diagnosed with certain hospital-acquired conditions (such as surgical-site infections following particular procedures and catheter-associated urinary tract infections). Including fewer measures could be less confusing for both consumers and providers because their attention would be focused on a select set of quality measures. But it is unclear whether efforts to improve according to narrower measures will lead to changes that yield broader gains in care quality or will instead yield only improvements according to those measures.

Some experts have expressed concerns that current public reporting requirements are overly burdensome for providers and that the measures required by Medicare are not aligned with those required by private insurers.[88] To that end, CMS, NQF, America's Health Insurance Plans—a

[85] While some measures, such as those involving surgical complications, rely entirely on the experience of Medicare patients, other measures, such as those collected through HCAHPS surveys of patient experience and condition-specific process measures, include non-Medicare patients as well.

[86] The measure of average spending per beneficiary captures all Medicare Part A and Part B spending for the period extending from 3 days before the hospital admission through 30 days after discharge. The measure of spending is computed for patients with specified diagnoses; it is risk-adjusted and adjusted to remove the effects of geographic differences in providers' input prices.

[87] Edward C. Norton and others, *Moneyball in Medicare*, Working Paper 22371 (National Bureau of Economic Research, June 2016), www.nber.org/papers/w22371.

[88] Medicare Payment Advisory Commission, "Next Steps in Measuring Quality of Care in Medicare," in *Medicare and the Health Care Delivery System* (June 2015), www.medpac.gov/docs/default-source/reports/june-2015-report-to-the-congress-medicare-and-the-health-care-delivery-system.pdf?sfvrsn=0; and Lawrence P. Casalino and others,

trade group for health insurers—and several other stakeholders announced the establishment of the Core Quality Measure Collaborative in February 2016 with the goal of aligning core quality measures among commercial and government payers. That organization has developed seven core sets of quality measures that can be used by all payer types, and CMS has begun using measures from each of the core sets. Commercial payers will incorporate those measures as their contracts with providers are renewed.

Quality programs have adjusted the composition of measures they include over the years in response to expert opinion about the types of measures that are most valuable and as data on new measures have become available. For example, when Medicare's Hospital Value-Based Purchasing Program began in 2013, 70 percent of a hospital's score was based on process measures and 30 percent was based on patient experience measures. Since then, new measures have been added and others have been dropped, and the emphasis on process measures has declined significantly. In addition, CMS changed the focus of the program substantially in 2015 when it added a measure of average spending per beneficiary. In 2016, outcome measures received a weight of 40 percent, patient experience measures and the spending measure each received a weight of 25 percent, and process measures received a weight of only 10 percent.[89] Similarly, Hospital Compare initially reported on a set of process measures, later adding outcome measures and patient experience measures. Medicare's physician quality reporting system also largely focused on process measures at first, but it has since started to incorporate patient experience measures and various types of outcome measures.

Risk Adjustment. Because quality scores often depend on both the performance of providers and the characteristics of their patients, a risk-adjustment mechanism is an essential feature of many public reporting programs and P4P programs. Outcome-based measures in particular are likely to vary with the characteristics and health status of patients, and if those factors are not adequately accounted for, the measures could give a misleading picture of the quality of care delivered by different providers—and P4P programs could unfairly penalize some providers and unfairly reward others. As discussed above, however, existing risk-adjustment methods have limitations, so even with those adjustments, the resulting quality measures would be somewhat noisy signals of true quality and could also create incentives for providers to avoid high-risk patients or to seek low-risk patients to boost their quality scores.[90]

"U.S. Physician Practices Spend More Than $15.4 Billion Annually to Report Quality Measures," *Health Affairs*, vol. 35, no. 3 (March 2016), pp. 401–406, https://dx.doi.org/10.1377/hlthaff.2015.1258.

[89] Lemeneh Tefera, "Hospital Value-Based Purchasing (HVBP) Program" (Centers for Medicare & Medicaid Services), https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/HVBP/2013-2015-HVBP-Presentation-.pptx.

[90] David Dranove and others, "Is More Information Better? The Effects of 'Report Cards' on Health Care Providers," *Journal of Political Economy*, vol. 111, no. 3 (June 2003), pp. 555–588, http://dx.doi.org/10.1086/374180; Joshua Burack and others, "Public Reporting of Surgical Mortality: A Survey of New York State Cardiothoracic Surgeons," *The Annals of Thoracic Surgery*, vol. 68, no. 4 (October 1999), pp. 1195–1202, http://dx.doi.org/10.1016/S0003-4975(99)00907-8; Eric C. Schneider and Arnold M. Epstein,

The factors that risk-adjustment systems account for typically vary with the measure. Most outcome and spending measures are risk-adjusted to account for age, sex, diagnosis, and comorbidities.[91] However, Medicare does not risk-adjust its measures of hospital-acquired conditions because the criteria for selecting conditions include a requirement that each condition be preventable by providers, regardless of the characteristics of the patient.[92] Likewise, Medicare uses specific criteria for determining whether particular treatments are appropriate for individual patients and does not adjust process measures because those measures are applied only to the set of patients for whom the treatments are deemed appropriate.[93] Medicare does adjust patient experience measures, but not for diagnoses or comorbidities. Instead, those measures are adjusted on the basis of age, education level, language spoken at home, and service line (including surgical, medical, and maternity care).[94] As discussed previously, there is debate

---

"Influence of Cardiac-Surgery Performance Reports on Referral Practices and Access to Care," *The New England Journal of Medicine*, vol. 335, no. 4 (July 1996), pp. 251–256, http://dx.doi.org/10.1056/NEJM199607253350406; Andrew M. Ryan and others, "The Relationship Between Medicare's Process of Care Quality Measures and Mortality," *Inquiry*, vol. 46, no. 3 (August 2009), pp. 274–290, http://dx.doi.org/10.5034/inquiryjrnl_46.03.274; Grace M. Lee and others, "Effect of Nonpayment for Preventable Infections in U.S. Hospitals," *The New England Journal of Medicine*, vol. 367, no. 15 (October 2012), pp. 1428–1437, http://dx.doi.org/10.1056/NEJMsa1202419; and Samuel K. Peasah and others, "Medicare Non-Payment of Hospital-Acquired Infections: Infection Rates Three Years Post Implementation," *Medicare & Medicaid Research Review*, vol. 3, no. 3 (2013), pp. E1–E13, http://dx.doi.org/10.5600/mmrr.003.03.a08.

[91] See, for example, Centers for Medicare & Medicaid Services, "Measures of Complications, Deaths, & Hospital Returns" (accessed October 6, 2017), www.medicare.gov/hospitalcompare/Data/RCD-Overview.html; Centers for Medicare & Medicaid Services, "Readmissions Reduction Program (HRRP)" (November 21, 2017), www.cms.gov/medicare/medicare-fee-for-service-payment/acuteinpatientpps/readmissions-reduction-program.html; Centers for Medicare & Medicaid Services, "Payment Measures" (accessed October 6, 2017), www.medicare.gov/hospitalcompare/Data/Payment-measures.html; Ashish Jha and others, "The Long-Term Effect of Premier Pay for Performance on Patient Outcomes," *The New England Journal of Medicine*, vol. 366, no. 17 (April 2012), pp. 1606–1615, http://dx.doi.org/10.1056/NEJMsa1112351;
Sule Calikoglu, Robert Murray, and Dianne Feeney, "Hospital Pay-For-Performance Programs in Maryland Produced Strong Results, Including Reduced Hospital-Acquired Conditions," *Health Affairs*, vol. 31, no.12 (December 2012), pp. 2649–2658, http://dx.doi.org/10.1377/hlthaff.2012.0357; Centers for Medicare & Medicaid Services, "Program Guidance & Specifications" (November 13, 2017), www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/sharedsavingsprogram/program-guidance-and-specifications.html; and Centers for Medicare & Medicaid Services, *2015 QRUR and the 2017 Value Modifier: Risk Adjustment* (March 2017), https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/PhysicianFeedbackProgram/Downloads/2015-RiskAdj-FactSheet.pdf. Nursing home quality measures use nursing home assessment data rather than claims data and therefore use different information, such as functional status, to adjust outcome measures. For more information on quality measurement in nursing homes, see Centers for Medicare & Medicaid Services, *MDS 3.0 Quality Measures: User's Manual* (April 2017), https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/NursingHomeQualityInits/Downloads/MDS-30-QM-Users-Manual-V11-Final.pdf.

[92] Sule Calikoglu, Robert Murray, and Dianne Feeney, "Hospital Pay-For-Performance Programs in Maryland Produced Strong Results, Including Reduced Hospital-Acquired Conditions," *Health Affairs*, vol. 31, no.12 (December 2012), pp. 2649–2658, http://dx.doi.org/10.1377/hlthaff.2012.0357.

[93] See, for example, Centers for Medicare & Medicaid Services, "Timely & Effective Care Measures" (accessed October 6, 2017), www.medicare.gov/hospitalcompare/Data/Measures.html.

[94] See Hospital Consumer Assessment of Healthcare Providers and Systems, "Mode & Patient-Mix Adjustment" (accessed on October 6, 2017), www.hcahpsonline.org/modeadjustment.aspx.

about whether certain quality measures should also be adjusted to account for differences in patients' socioeconomic status. Most recently, the 21st Century Cures Act, which was passed in December 2016, requires CMS to stratify hospitals in the Hospital Readmissions Reduction Program on the basis of the share of their Medicare patients who are dually eligible—that is, enrolled in both Medicare and Medicaid.

Presentation of Measures to Consumers. Another important consideration in designing quality measurement initiatives is how to present the information to consumers. There has been substantial research into the design of quality reports in order to determine what types of presentation strategies are most easily understood by consumers. Given the complexity of the information contained in such reports, one primary concern is to avoid information overload, which results in consumers' needing more time to process the information included or possibly ignoring the information altogether.[95] Research has found that consumers understand information best when it is presented simply, through strategies such as symbols (for example, star ratings), summary measures, rank ordering or tiering, and consistency among measures so that either high marks or low marks consistently indicate better performance.[96]

Public reporting programs have employed a variety of those strategies in order to better engage consumers. For example, Hospital Compare began using a composite star rating for its patient experience measures in 2015 and added an overall composite star rating in 2016—though some researchers have identified challenges with that measure, as discussed above. Nursing Home Compare and Home Health Compare also report an overall composite star rating in addition to the underlying quality measures. Those composites may be simpler for patients to understand, but they still require that providers collect and report data for each individual quality measure. In addition, patients still may need to "drill down" to the underlying measures to compare how hospitals perform in select areas or for treating select conditions. Furthermore, while Medicare's reporting programs do not structure their quality measures so that high (or low) marks consistently indicate better performance, they do specify whether higher or lower scores indicate better performance on individual measures.

Incentive Structure. As discussed above, there are two broad approaches used: public reporting alone, which can indirectly provide incentives to improve quality scores; and P4P programs, which combine public reporting with financial incentives tied to quality. This section discusses the nature of those direct financial incentives, their size, the level of the provider organizations

---

[95] Zoe Hildon, Dominique Allwood, and Nick Black, "Impact of Format and Content of Visual Display of Data on Comprehension, Choice and Preference: A Systematic Review," *International Journal for Quality in Health Care*, vol. 24, no. 1 (December 2011), pp. 55–64, http://dx.doi.org/10.1093/intqhc/mzr072.

[96] Agency for Healthcare Quality and Research, *Best Practices in Public Reporting No. 1: How to Effectively Present Health Care Performance Data to Consumers* (prepared by the Center for Health Improvement, June 2010), http://archive.ahrq.gov/professionals/quality-patient-safety/quality-resources/tools/pubrptguide1/pubrptguide1.pdf; and Aligning Forces for Quality, *How to Display Comparative Information That People Can Understand and Use* (Robert Wood Johnson Foundation, 2010), www.rwjf.org/content/dam/farm/reports/reports/2010/rwjf69342.

that are targeted, the timing of the incentives, and requirements regarding providers' participation.

Nature of the Incentives. Financial incentives may be bonuses, penalties, or a combination of both. In applying those incentives, programs can evaluate providers' performance against a specified threshold (sometimes called an absolute benchmark) or against the performance of other providers (a relative benchmark). A second way incentive payments may vary is whether they are based on the level of the scores, thus measuring *achievement*, or instead on the amount of *improvement* in the scores. A third choice is whether to structure the incentives as bonuses for good performance or penalties for subpar performance. Researchers have found that people take greater efforts to avoid potential losses than they do to obtain potential gains of a similar amount—which suggests that providers may do more to improve quality to avoid penalties than they would to receive bonuses.[97]

Each approach offers advantages and disadvantages. Rewarding performance relative to a threshold ensures that providers are rewarded for achieving a certain benchmark of quality. However, those thresholds may be too high a bar for low performers, who then would have little incentive to improve. Rewarding improvement creates stronger incentives for lower-quality providers, but providers with very high quality scores and limited room for further improvement might view that type of system as unfair. Recognizing those trade-offs, some programs use a combination of approaches in which providers are evaluated partly on the basis of improvement and partly on reaching a certain benchmark for quality.

Medicare's Hospital Value-Based Purchasing Program is an example of a program that uses both bonuses and penalties and evaluates hospitals both on how their quality scores compare with those of their peers and on the improvement in their quality scores. Under that program, CMS withholds a portion of Medicare payments to hospitals to form a bonus pool. Those withheld payments are then redistributed to hospitals to reflect their quality performance. Worse-performing hospitals earn bonuses that are smaller than the withheld amount and thus end up with lower total payments, whereas better-performing hospitals earn bonuses that are larger than the withheld amount and come out ahead.

The Premier Hospital Quality Incentive Demonstration is another example of a program that used both achievement and improvement to determine hospitals' bonuses. The Premier demonstration initially rewarded hospitals on the basis of absolute benchmarks but adjusted its incentive design in a second phase to also reward hospitals that improved their quality performance. Research has found, however, that while hospitals participating in that demonstration improved quality more quickly than nonparticipants in the first few years of the

[97] Ateev Mehrotra, Melony E. S. Sorbero, and Cheryl L. Damberg, "Using the Lessons of Behavioral Economics to Design More Effective Pay-for-Performance Programs," *The American Journal of Managed Care*, vol. 16, no. 7 (July 2010), pp. 497–503, www.ncbi.nlm.nih.gov/pmc/articles/PMC3303871/.

demonstration, nonparticipating hospitals caught up in subsequent years. Five years from the start of the demonstration, participating hospitals achieved no better quality improvement than nonparticipating hospitals did.[98]

Medicare's Hospital Readmissions Reduction Program, Hospital-Acquired Condition Reduction Program, and End-Stage Renal Disease Quality Incentive Program are examples of programs that use penalties alone to incentivize providers. The first two programs evaluate hospitals on the basis of their performance relative to other hospitals', while the third program compares performance to prespecified thresholds. The Hospital Readmissions Reduction Program penalizes hospitals whose readmission rates for selected conditions are above the national average rates for a prior three-year period, and the Hospital-Acquired Condition Reduction Program penalizes hospitals in the lowest-performing quartile nationally.[99] Some analysts have criticized those incentive structures because the total amount of the penalties levied on hospitals would not decline even if all hospitals reduced their rates of readmissions and hospital-acquired conditions.[100] Proponents of the program, however, might argue that it maintains continuous pressure on hospitals to reduce those rates, whereas a program that was based on an absolute benchmark would not have that effect.

Size of the Incentives. Another consideration for financial incentives is their size. Policymakers and researchers have not reached a consensus on the optimal incentive size. Some have expressed concerns that incentives in federal programs are too small to affect providers' behavior.[101] Others have argued that financial incentives in P4P programs can actually lead to poorer performance in fields such as health care, which involve cognitively complex and intrinsically rewarding tasks, and that the reduction in performance might be greatest when the incentives are large.[102] At root, these debates reflect uncertainty about how much it may cost to

[98] Rachel M. Werner and others, "The Effect of Pay-for-Performance in Hospitals: Lessons for Quality Improvement," *Health Affairs*, vol. 30, no. 4 (April 2011), pp. 690–698, http://dx.doi.org/10.1377/hlthaff.2010.1277.

[99] In addition to this penalty, Medicare no longer makes additional payments to hospitals for a select set of hospital-acquired conditions that were not present upon a patient's admission to the hospital. CMS refers to this program as the "Hospital-Acquired Condition Present on Admission Indicator" program. Previously, those conditions qualified hospitals for an additional payment for treatments of patients' complications.

[100] For example, see Medicare Payment Advisory Commission, "Refining the Hospital Readmissions Reduction Program," in *Medicare and the Health Care Delivery System* (June 2013), www.medpac.gov/docs/default-source/reports/jun13_ch04.pdf?sfvrsn=0.

[101] Rachel M. Werner and R. Adams Dudley, "Medicare's New Hospital Value-Based Purchasing Program Is Likely to Have Only a Small Impact on Hospital Payments," *Health Affairs*, vol. 31, no. 9 (September 2012), pp. 1932–1940, http://dx.doi.org/10.1377/hlthaff.2011.0990.

[102] Steffie Woolhandler and Dan Ariely, "Will Pay for Performance Backfire? Insights from Behavioral Economics," *Health Affairs Blog* (October 11, 2012), http://healthaffairs.org/blog/2012/10/11/will-pay-for-performance-backfire-insights-from-behavioral-economics.

improve care quality and how much those improvements are worth in terms of health gains or potential savings.

Most P4P programs have relatively small incentives as a percentage of total payments. For example, the incentives in Medicare's Hospital Value-Based Purchasing Program range between a 2.0 percent reduction and a 3.5 percent increase in payments in 2017. In that year, Medicare's Hospital Readmissions Reduction Program assigned penalties of up to 3.0 percent, and Medicare's Hospital-Acquired Condition Reduction program assigned penalties of 1.0 percent for the worst-performing 25 percent of hospitals.[103] While each of those incentives might seem small on its own, their combined effect ranges from a 6.0 percent reduction in payments to a 3.5 percent increase. Moreover, the incentive payments are larger as a share of providers' operating margins and thus may have important effects on whether providers turn a profit or lose money.[104]

So far, hospitals' payment adjustments within each program have been minimal in most cases. In 2017, under the Hospital Value-Based Purchasing Program, 54 percent of participating hospitals had adjustments within 0.5 percent of their original payments, and 84 percent of hospitals had adjustments within 1.0 percent of their original payments. In the same year, 40 percent of hospitals were penalized by 0.5 percent or more under Medicare's Hospital Readmissions Reduction Program, and 20 percent of hospitals were penalized by 1 percent or more.[105] However, the cumulative effect of those financial incentives was more substantial for some hospitals and has been found to vary with hospitals' characteristics. For example, teaching hospitals and those with larger bed counts were particularly likely to be penalized under federal programs in 2015: The cumulative payment adjustment for the three programs described above averaged -0.5 percent among all hospitals but averaged -0.9 percent for teaching hospitals and -0.8 percent for hospitals with more than 400 beds.[106]

Organizational Level of Targeted Providers. Another design consideration for quality initiatives is the provider level that they should target. There is an ongoing debate about whether programs should target individual physicians—so that each physician is evaluated on the basis of his or her

[103] Medicare Payment Advisory Commission, "Hospital Inpatient and Outpatient Services," in *Medicare Payment Policy* (March 2017), www.medpac.gov/docs/default-source/reports/mar17_entirereport224610adfa9c665e80adff00009edf9c.pdf?sfvrsn=0.

[104] For example, CBO found that average hospital margins were 6.0 percent in 2011. See Tamara Hayford, Lyle Nelson, and Alexia Diorio, *Projecting Hospitals' Profit Margins Under Several Illustrative Scenarios*, Working Paper 2016-04 (Congressional Budget Office, September 2016), www.cbo.gov/publication/51919.

[105] Centers for Medicare & Medicaid Services, "FY 2017 Final Rule and Correction Notice Tables" (accessed October 6, 2017), Tables 15, 16A, and 16B, www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS/FY2017-IPPS-Final-Rule-Home-Page-Items/FY2017-IPPS-Final-Rule-Tables.html.

[106] Charles N. Kahn and others, "Assessing Medicare's Hospital Pay-for-Performance Programs and Whether They Are Achieving Their Goals," *Health Affairs*, vol. 34, no. 8 (August 2015), pp. 1281–1288, http://dx.doi.org/10.1377/hlthaff.2015.0158.

own performance—or the group or practice level.[107] Some arguments in favor of targeting incentives at the practice level are that systemic changes are made at the practice level, there is greater reliability of quality measurement because of larger sample sizes, and practice-level incentives might facilitate greater cooperation. In addition, the practice is the natural unit for many payment and contracting decisions by health plans. However, practice-level incentives might weaken individual providers' incentives to improve quality, and it may be difficult to attribute a practice-level quality measure to any individual provider. Individual-level incentives may be particularly important in areas in which individual physicians' decisions contribute to performance, such as the provision of preventive services. Programs can also include a blend of practice-level and physician-level targeting.[108] Historically, Medicare has evaluated both ACOs and physicians at the practice level, but it will begin evaluating individual physicians under the Quality Payment Program as early as 2017.

Timing of the Incentives. Whether the incentives are reports of quality achievement under public reporting or financial bonuses and penalties under P4P programs, the timing of the incentives relative to the performance period may influence the effectiveness of the quality initiative. Longer lag times between performance and the resulting incentive payments may reduce the incentive for providers to make improvements or correct deficiencies. However, some measures—such as condition-specific readmission rates—are based on infrequent events and require multiple years of data to have an adequate sample size for smaller providers. Data collection may play an additional role in the length of these lags. If there is a delay in adjudicating a subset of the claims, claims-based data may only be available with a lag. While medical records could be abstracted in a more timely fashion, abstraction from paper records is nevertheless a time-consuming process. Electronic health records may ease that administrative burden in some cases, but the use of different software systems by different providers may complicate the collection of those data.

Because of the time involved in adjudicating claims, most Medicare quality initiatives incorporate a lag of at least 18 months—and in some cases up to four years. For example, the penalty for Medicare's Hospital Readmissions Reduction Program is based on the most recent three years of available data—so the 2016 penalty, for example, is based on readmissions

[107] The American Medical Association found that 17 percent of physicians were in a solo practice in 2014. See American Medical Association, "AMA Study Finds Majority of Physicians Still Work in Small Practices" (press release, July 8, 2015), https://www.ama-assn.org/content/new-ama-study-reveals-majority-americas-physicians-still-work-small-practices.

[108] Meredith B. Rosenthal and R. Adams Dudley, "Pay for Performance: Will the Latest Payment Trend Improve Care?" *Journal of the American Medical Association*, vol. 297, no. 7 (February 2007), pp. 740–744, http://dx.doi.org/10.1001/jama.297.7.740; and Jessica Greene, Ellen Kurtzman, Judith Hibbard, and Valerie Overton, "Working Under a Clinic-Level Quality Incentive: Primary Care Clinicians' Perceptions," *Annals of Family Medicine*, vol. 13, no. 3 (May/June 2015), pp. 235–241, http://dx.doi.org/10.1370/afm.1779.

occurring from 2012 through 2014.[109] Similarly, the Physician Value-Based Modifier Program, which assigns bonuses or penalties to provider groups on the basis of their performance on various quality and cost measures, operates with a two-year lag—the 2016 adjustment is based on 2014 data—whereas the Hospital Value-Based Purchasing Program assigns payment adjustments on the basis of the prior year's quality levels with a six-month lag.[110] Medicare's Home Health Value-Based Purchasing Demonstration, which will begin adjusting payment rates for home health agencies in select states in 2018, will also assign bonuses and penalties on the basis of performance from two years earlier. One notable exception is Medicare's requirement that certain conditions be documented as "present on admission" in order for Medicare to reimburse hospitals for their treatment. That program has a minimal lag because it is administered at the point of payment: CMS does not reimburse hospitals for the additional costs associated with treating certain hospital-acquired conditions if they were not present upon admission.

Program Participation Rules. Another design feature of quality initiatives is the rules governing program participation. Participation by providers is mandatory in some programs and voluntary in others, although some voluntary programs reduce payments to providers that do not participate. For example, Medicare requires all acute care hospitals paid under the inpatient prospective payment system to participate in its Hospital Value-Based Purchasing Program and Hospital-Acquired Condition Reduction Program. In the Hospital Value-Based Purchasing Program, CMS automatically withholds that year's contribution to the bonus pool and calculates the quality score for all hospitals that report data on a core set of quality measures. Although hospitals are not required to submit data on the core set of quality measures to Medicare, they must do so in order to receive a full payment update from Medicare; otherwise, they face a 2 percent payment reduction.[111] Because of the potential for that payment reduction, virtually all hospitals participate. In the Hospital-Acquired Condition Reduction Program, failure by nonexempt hospitals to report data on any of the required quality measures that are not based on claims has a direct impact on the hospital's program score, and it increases the likelihood of being penalized.[112]

---

[109] Assessment periods run from July through June, so the assessment period for the 2016 fiscal year penalty runs from July 2011 through June 2014.

[110] Because of sample size considerations, mortality and the patient safety composite measure are measured for nearly two years. For more information, see Centers for Medicare & Medicaid Services, "Hospital Value-Based Purchasing" (September 2015), www.cms.gov/Outreach-and-Education/Medicare-Learning-Network-MLN/MLNProducts/downloads/Hospital_VBPurchasing_Fact_Sheet_ICN907664.pdf, and "Computation of the 2016 Value Modifier" (September 2015), www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/PhysicianFeedbackProgram/Downloads/2016-VM-Fact-Sheet.pdf.

[111] Centers for Medicare & Medicaid Services, "Hospital Inpatient Quality Reporting Program" (September 19, 2017), www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/HospitalRHQDAPU.html.

[112] See Quality Net, *Hospital-Acquired Condition Reduction Program: Frequently Asked Questions, Fiscal Year 2018,*

Physicians and other eligible professionals may voluntarily report on quality measures and receive bonus payments under the Physician Quality Reporting System. However, starting in 2017, CMS will reduce payments to physicians who did not report data in 2015 by 2 percent.[113] Hospitals, physicians, and other health care providers may also choose whether or not to participate in CMS's ACO programs (the Medicare Shared Savings Program, the Advanced Payment ACO Model, the ACO Investment Model, the Next Generation ACO Model, and the Pioneer ACO Model).

With the passage of the Medicare Access and CHIP Reauthorization Act of 2015, however, physicians will be required to participate in one of two quality-based payment tracks under the Quality Payment Program: the Merit-Based Incentive Payment System and alternative payment models. The Merit-Based Incentive Payment System will adjust physicians' payments on the basis of performance in four areas: quality, resource use, clinical practice improvement, and meaningful use of electronic health records. Alternative payment models include ACOs, the Comprehensive Primary Care Plus Initiative, and other types of demonstrations. Those alternative models place additional requirements on practices regarding the use of quality measurement, the use of electronic health records, and the assumption of financial risk for patients' spending patterns.

4. Future Directions

Measuring the quality of providers has become increasingly common in recent years, but measuring it accurately and reporting those results remains a challenging task. Although providers have made some progress in improving quality, those improvements have been slow in coming, and deficiencies in quality still persist. As a result, some researchers have called for the development of better measures and a more effective application of existing ones.

Development of Better Measures

Researchers are trying to address the limitations of existing quality measures by developing new ones. Those efforts fall into two broad categories: improving existing types of measures (for example, process, outcome, and composite measures) and developing new types of measures that

www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier3&cid=1228774298670.

[113] Centers for Medicare & Medicaid Services, "2015–2017 Physician Quality Reporting System (PQRS) Timeline" (January 2016), www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/PQRS/Downloads/2015-17_CMS_PQRS_Timeline.pdf. The Medicare Access and CHIP Reauthorization Act of 2015 phases out the individual Physician Quality Reporting System. Beginning in 2019, CMS will instead evaluate providers under the Merit-Based Incentive Payment System (MIPS). It includes four categories: quality, resource use, clinical practice improvement activities, and EHR meaningful use. Each professional's MIPS-eligible professional will receive a composite score that will be measured against a performance threshold and will result in an upward, downward, or neutral payment adjustment.

encompass the continuum of care, focus more on systems of care, and seek to measure the health of populations rather than individual patients.

Improvement of Existing Types of Measures. Measures based on processes of care have historically focused on underuse. More recently, however, researchers have focused on developing measures of overuse as a way to increase the efficiency of the health care system and reduce unnecessary testing and treatment.[114] Overuse measures are aimed at areas of care in which there is strong evidence that a given service will not improve patient outcomes or is potentially harmful. However, a major concern with overuse measures is that they might lead to the underuse of appropriate treatments and tests because of the difficulty of correctly identifying instances of overuse.[115]

Additionally, some researchers have recommended refining existing outcome measures and the evaluation of performance on those measures. For example, researchers have suggested that CMS revise its Hospital Readmissions Reduction Program to evaluate hospitals' performance against a fixed target rather than relative to the performance of other hospitals, so that aggregate penalties would decline when readmission rates improved throughout the sector. Researchers have also recommended the adoption of an all-condition readmissions measure because small sample sizes may cause existing condition-specific measures to be more volatile for some providers. Additionally, some analysts have proposed that CMS account for patients' socioeconomic status in its readmissions measures, reflecting the concern that safety-net hospitals are penalized because they treat a large proportion of low-income patients.[116] In 2019, in accordance with the 21st Century Cures Act, CMS will begin comparing hospitals with peer institutions with a similar proportion of Medicare beneficiaries who are dually eligible with full Medicaid benefits.

Patient-reported outcomes (PROs) are another type of measure that could play a larger role in assessing the quality of patients who are recovering from an acute illness or living with a chronic condition. Under MACRA, CMS has identified the development of PRO measures as a priority area for the agency.[117] Although PROs have been used in clinical studies, they have not been

---

[114] Mark R. Chassin, "Improving the Quality of Health Care: What's Taking So Long?" *Health Affairs*, vol. 32, no. 10 (October 2013), pp. 1761–1765, http://dx.doi.org/10.1377/hlthaff.2013.0809; and Deborah Korenstein and others, "Overuse of Health Care Services in the United States: An Understudied Problem," *Archive of Internal Medicine*, vol. 172, no. 2 (January 2012), pp. 171–178, http://dx.doi.org/10.1001/archinternmed.2011.772.

[115] Jason S. Mathias and David W. Baker, "Developing Quality Measures to Address Overuse," *Journal of the American Medical Association*, vol. 309, no. 18 (May 2013), pp. 1897–1898, http://dx.doi.org/10.1001/jama.2013.3588.

[116] Andrew S. Boozary, Joseph Manchin III, and Roger F. Wicker, "The Medicare Hospital Readmissions Reduction Program: Time for Reform," *Journal of the American Medical Association*, vol. 314, no. 4 (July 2015), pp. 347–348, http://dx.doi.org/10.1001/jama.2015.6507.

[117] Centers for Medicare & Medicaid Services, "CMS Quality Measurement Development Plan: Supporting the Transition to Merit-Based Incentive Payment System (MIPS) and Alternative Payment Models (APMs)" (May

routinely used in clinical practice because of logistical and measurement challenges that hinder widespread implementation.[118] PROs may also need to evolve further before they are incorporated in value-based payment programs.

Finally, researchers are refining existing composite measures and developing new ones to more accurately measure quality, particularly at the physician level. Several studies have found that composite measures of common inpatient conditions and surgical quality are more effective than current measures at predicting mortality and helping consumers distinguish between high and low performers.[119] But how well a composite measure performs depends in part on the underlying individual measures that are used to construct it and how they are weighted.

Development of New Types of Measures. Researchers have also proposed developing new types of measures that assess quality across the continuum of care instead of focusing on the performance of individual providers or specific settings (for instance, hospitals, nursing homes, or home health). Poor quality is often viewed as the failure of an individual provider rather than the system of care. This view may reinforce the fragmentation of the health care system.[120] But there is a growing recognition that health care is generally delivered across multiple settings by teams of specialists, nurses, and other providers. Therefore, evaluating quality across the care continuum would emphasize the health care team, and it could overcome some of the technical challenges of measuring individual providers—small sample size and problems of attribution.[121] However, being able to attribute specific care processes to individual team members may be difficult. This approach to quality measurement would also still require that payers monitor individual providers for poor performance.[122]

2016), https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/Final-MDP.pdf.

[118] Danielle C. Lavallee and others, "Incorporating Patient-Reported Outcomes Into Health Care to Engage Patients and Enhance Care," *Health Affairs*, vol. 35, no. 4 (April 2016), pp. 575–582, http://dx.doi.org/10.1377/hlthaff.2015.1362.

[119] Lena M. Chen and others, "Composite Quality Measures for Common Inpatient Medical Conditions," *Medical Care*, vol. 51, no. 9, pp. 832–837 (September 2013), http://dx.doi.org/10.1097/MLR.0b013e31829fa92a; and Justin B. Dimick and others, "Composite Measures for Rating Hospital Quality With Major Surgery," *Health Services Research*, vol. 47, no. 5 (October 2012), pp. 1861–1879, http://dx.doi.org/10.1111/j.1475-6773.2012.01407.x.

[120] Elliott S. Fisher and Stephen M. Shortell, "Accountable Care Organizations: Accountable for What, to Whom, and How," *Journal of the American Medical Association*, vol. 304, no. 15 (October 2010), http://dx.doi.org/10.1001/jama.2010.1513.

[121] Elliott S. Fisher and Stephen M. Shortell, "Accountable Care Organizations: Accountable for What, to Whom, and How," *Journal of the American Medical Association*, vol. 304, no. 15 (October 2010), http://dx.doi.org/10.1001/jama.2010.1513; and Robert A. Berenson, Peter J. Pronovost, and Harlan M. Krumholz, "Achieving the Potential of Health Care Performance Measures" (Robert Wood Johnson Foundation and Urban Institute, May 2013), www.rwjf.org/content/dam/farm/reports/reports/2013/rwjf406195.

[122] Robert A. Berenson, Peter J. Pronovost, and Harlan M. Krumholz, "Achieving the Potential of Health Care Performance Measures" (Robert Wood Johnson Foundation and Urban Institute, May 2013), www.rwjf.org/content/dam/farm/reports/reports/2013/rwjf406195.

In accordance with the above framework, CMS is starting to reorient measures to span multiple settings of care. Patients' experiences and outcomes relating to transitions between settings (such as discharges from hospitals to skilled nursing facilities or to patients' homes) are of particular interest.[123]

Researchers have also been placing increased emphasis on developing population-based outcome measures. Such measures evaluate the quality of care that is delivered by health care systems to a population in a local area or to a set of people enrolled in the same program. For example, the Medicare Payment Advisory Commission (MedPAC) has proposed measuring potentially preventable hospital admissions, potentially preventable emergency visits, healthy days at home, and patient experience measures at the population level. MedPAC has also proposed that CMS could measure the performance of ACOs and Medicare Advantage (MA) plans for their enrollees at the local market–area level against local market–area quality benchmarks. The difference between that measurement program and others discussed above is that the quality measures would reflect systemic quality by payment model rather than provider-specific quality.[124] Although CMS already assesses the quality of MA plans through its star rating program, MA plans' quality are not currently rated (and their payments are not adjusted) in relation to the performance of ACOs or providers within a fee-for-service setting.

A limitation of the above approach, however, is that it would not be appropriate to make quality-based payment adjustments for physicians' treatment of patients who are *not* participating in an ACO or MA plan, because there is no single entity that is responsible for the full continuum of care for that population. Thus, Medicare would still need to use provider-level measures to evaluate providers whose patients are generally not enrolled in those programs.[125]

More Effective Application of Measures

In addition to the efforts by payers to improve existing measures and develop new ones, researchers have recommended ways for payers to apply quality measures more effectively. For instance, researchers have recommended that payers modify existing programs to be more user-

[123] Patrick H. Conway, Farzad Mostashari, and Carolyn Clancy, "The Future of Quality Measurement for Improvement and Accountability," *Journal of the American Medical Association*, vol. 309, no. 21 (June 2013), http://dx.doi.org/10.1001/jama.2013.4929.

[124] Medicare Payment Advisory Commission, "Measuring Quality of Care in Medicare," in *Medicare and the Health Care Delivery System* (June 2014), pp. 39–56, www.medpac.gov/docs/default-source/reports/chapter-3-measuring-quality-of-care-in-medicare-june-2014-report-.pdf?sfvrsn=0, and "Next Steps in Measuring Quality of Care in Medicare," in *Medicare and the Health Care Delivery System* (June 2015), www.medpac.gov/docs/default-source/reports/june-2015-report-to-the-congress-medicare-and-the-health-care-delivery-system.pdf?sfvrsn=0.

[125] Medicare Payment Advisory Commission, "Measuring Quality of Care in Medicare," in *Medicare and the Health Care Delivery System* (June 2014), pp. 39–56, www.medpac.gov/docs/default-source/reports/chapter-3-measuring-quality-of-care-in-medicare-june-2014-report-.pdf?sfvrsn=0, and "Next Steps in Measuring Quality of Care in Medicare," in *Medicare and the Health Care Delivery System* (June 2015), www.medpac.gov/docs/default-source/reports/june-2015-report-to-the-congress-medicare-and-the-health-care-delivery-system.pdf?sfvrsn=0.

friendly and relevant to consumers, and they have also suggested ways to redesign P4P programs to increase their effectiveness.

A recent report by the Government Accountability Office (GAO) found that CMS's public reporting programs provide limited information about providers' quality that is relevant and understandable to consumers.[126] According to the GAO report, those programs have not been able to meaningfully distinguish performance among providers, in part because of confusing layouts, data gaps, and lack of customizable information. GAO has recommended that CMS improve the organization of quality reports to better identify the highest-performing providers, customize information so that it is more relevant for consumers, and ensure that performance measures are consumer-centric.[127]

Partly in response to the GAO report, CMS recently added an overall five-star rating system for hospitals in Hospital Compare that supplements a star rating system based on patient experience scores.[128] However, as previously noted, some analysts have raised concerns about the overall star rating system in its current form.[129] CMS has also continued to update Physician Compare by adding patient experience measures in 2014 and clinician-level data in 2015. In late 2017, CMS plans to release star ratings that describe providers' performance on specific measures.[130]

Researchers have also called for redesigning P4P programs. Those recommendations include increasing the frequency of financial incentive payments, restructuring payment incentives to be more simplified and transparent, and requiring providers to make payments for poor performance instead of withholding or paying sums up front. The act of making a payment after performance

---

[126] Government Accountability Office, *Health Care Transparency: Actions Needed to Improve Cost and Quality Information for Consumers*, GAO-15-11 (October 2014), www.gao.gov/assets/670/666572.pdf.

[127] Ibid.

[128] Centers for Medicare & Medicaid Services, "First Release of the Overall Hospital Quality Star Rating on Hospital Compare" (press release, July 27, 2016), www.cms.gov/Newsroom/MediaReleaseDatabase/Fact-sheets/2016-Fact-sheets-items/2016-07-27.html, and "Hospital Compare Star Ratings Fact Sheet" (press release, April 16, 2015), www.cms.gov/Newsroom/MediaReleaseDatabase/Fact-sheets/2015-Fact-sheets-items/2015-04-16.html.

[129] Francis J. Crosson, Medicare Payment Advisory Commission, letter to the Honorable Andrew Slavitt, Acting Administrator, Centers for Medicare & Medicaid Services (September 22, 2016); Karl Y. Bilimoria and Cynthia Barnard, "The New CMS Hospital Quality Star Ratings: The Stars Are Not Aligned," *Journal of the American Medical Association*, vol. 316, no. 17 (November 2016), pp. 1761–1762, http://dx.doi.org/10.1001/jama.2016.13679; and Susan Xu and Atul Grover, "CMS' Hospital Quality Star Ratings Fail to Pass the Common Sense Test," *Health Affairs Blog* (November 14, 2016), http://healthaffairs.org/blog/2016/11/14/cms-hospital-quality-star-ratings-fail-to-pass-the-common-sense-test/.

[130] Centers for Medicare & Medicaid Services. "Quality Data and Physician Compare" (June 15, 2016), https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/physician-compare-initiative/Quality-Data-and-Physician-Compare-.html; and Alesia Hovatter, Denise St. Clair, and Allison Newsom, "Physician Compare: What You Need to Know" (Medicare Learning Network and Centers for Medicare & Medicaid Services, September 2017), https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/physician-compare-initiative/Downloads/Physician_Compare_National_Provider_Call-.pdf.

has been evaluated might induce a stronger behavioral response by providers than would seeing the payment not show up at all. Also, giving providers more real-time feedback and using tiered thresholds of performance rather than one absolute threshold might increase the effectiveness of P4P programs, especially if a provider's performance is far from the performance goal.[131] Another approach that researchers have suggested is applying financial incentives for providers to a subset of patients that are prospectively identified as being at the highest risk for poor outcomes.[132] That approach would send a stronger signal to providers about which patients need more attention to improve population health—so long as the method of identifying high-risk patients was effective. How all of these efforts to improve quality might play out in future programs, however, is difficult to predict.

[131] Ashish K. Jha, "Time to Get Serious About Pay for Performance," *Journal of the American Medical Association*, vol. 309, no. 4 (January 2013), pp. 347–348, http://dx.doi.org/10.1001/jama.2012.196646; Ateev Mehrotra, Melony E. S. Sorbero, and Cheryl L. Damberg, "Using the Lessons of Behavioral Economics to Design More Effective Pay-for-Performance Programs," *The American Journal of Managed Care*, vol. 16, no. 7 (July 2010), pp. 497–503, www.ncbi.nlm.nih.gov/pmc/articles/PMC3303871/; and Alliance for Health Reform, "Incentives 2.0: Is Paying for Performance Enough?" (October 2014), www.allhealth.org/briefingmaterials/TRANSCRIPT101414_1K.PDF.

[132] Aaron McKethan and Ashish K. Jha, "Designing Smarter Pay-for-Performance Programs," *Journal of the American Medical Association*, vol. 312, no. 24 (November 2014), pp. 2617–2618, http://dx.doi.org/10.1001/jama.2014.15398.

Tables

Table 1.
Public Reporting Programs Used by Medicare

Program Name	Type of Measures	Description
Hospital Compare	Structure, process, outcome, patient experience, payment/value calculations	Hospital Compare features a website that provides information on how well hospitals provide recommended care to patients. The program allows the comparison of performance measures related to heart attacks, heart failure, pneumonia, surgery, and other conditions. Hospital Compare recently added an overall five-star rating system.
Physician Compare	Process, patient safety, patient experience	Physician Compare provides information on physicians who participate in Medicare. The program began reporting process and outcome measures for group practices in February 2014 and has since added patient experience measures and clinician-level quality measures. The Centers for Medicare & Medicaid Services plans to introduce a composite star rating system by the end of 2017.
Nursing Home Compare	Structure, process, outcome, functional status	This program features a website that provides information on how well Medicare- and Medicaid-certified nursing homes are performing. The website includes a five-star rating system of nursing homes based on health inspections, quality measures, and staffing. In April 2016, Nursing Home Compare added measures related to outcomes of short-stay residents.
Home Health Compare	Process, outcome, patient experience	Home Health Compare features a website that provides information on the quality of care at Medicare-certified home health agencies. The program includes a five-star rating system for patient care and patient experiences.
Dialysis Facility Compare	Process, outcome, patient experience	This program compares the service offerings and quality of care that dialysis facilities provide. The Dialysis Facility Compare star ratings provide a summary measure of quality based on selected measures.

Note: This table excludes quality reporting programs for Medicare Advantage and Medicare Part D plans because they are outside the scope of this paper.

Table 2.
Medicare's Pay-for-Performance and Demonstration Programs

Program Name	Type of Measures	Description
Programs With Quality as an Objective Targeted at Hospitals		
Hospital Value-Based Purchasing Program	Process, outcome, patient experience, patient safety, efficiency	This program withholds a small portion of payments to hospitals and then redistributes them according to performance on quality measures. Redistributed payments are assigned to hospitals on the basis of overall performance and improvement.
Hospital Readmissions Reduction Program	Outcome (specifically, rate of 30-day readmissions)	The Hospital Readmissions Reduction Program requires the Medicare program to reduce the amount of payment to hospitals paid under the inpatient prospective payment system with excess readmissions.
Hospital-Acquired Condition Reduction Program	Outcome (specifically, rate of hospital-acquired conditions)	Under this program, hospitals in the worst-ranking quartile of quality measures for hospital-acquired conditions have their payments reduced by 1 percent.
Hospital-Acquired Condition Present-on-Admission (POA) Indicator program	Outcome (specifically, presence of hospital-acquired conditions following admission)	In the Hospital-Acquired Condition Present-on-Admission Indicator program, hospitals do not receive additional payment for cases in which the selected hospital-acquired condition was not present on admission.
Programs With Quality as an Objective Targeted at Other Types of Providers		
Medicare Access & CHIP Reauthorization Act (MACRA) Quality Payment Program	Process, outcome, patient experience, patient safety, care coordination, efficiency, population and community health	This Program established two paths for physicians: the Merit-Based Incentive Payment System (MIPS) and alternative payment models (APMs). The MIPS combines the Physician Quality Reporting System, the Value Modifier, and the Electronic Health Record incentive program into a single program that assigns bonuses or penalties to physicians on the basis of quality, resource use, clinical practice improvement, and meaningful use of electronic health records. Alternatively, physicians can participate in APMs that focus on paying providers for value, such as accountable care organizations, bundled payment models, or patient-centered medical homes.
Physician Value-Based Modifier Program	Effective clinical care, person and caregiver-centered experience and outcomes, population and community health, patient safety, communication and care coordination, efficiency and cost reduction, per capita costs	In this program, payments to physicians and physician groups under the Medicare Physician Fee Schedule are adjusted on the basis of the cost and quality of care delivered to Medicare fee-for-service beneficiaries.

(continued)

(continued)

Nursing Home Value-Based Purchasing Demonstration	Structure, outcome	In the Nursing Home Value-Based Purchasing Demonstration, Medicare assesses the performance of nursing homes on selected quality measures and makes incentive payments to those that perform the best or improve the most.
Home Health Value-Based Purchasing Model	Process, outcome, patient satisfaction	This demonstration ties home health payment to quality performance, and payments are adjusted up or down depending on quality achievement and improvement in relation to other facilities in the same size cohort and state.
End Stage Renal Disease (ESRD) Quality Incentive Program	Clinical, patient experience	The ESRD Quality Incentive Program links a portion of dialysis facility payments to performance on quality measures. The program reduces payments to facilities that do not meet or exceed performance standards.
Programs With Quality as a Constraint		
Accountable care organization (ACO) programs	Process, outcome, patient experience, patient safety	Under the Medicare ACO demonstration programs, bonus payments to providers based on any savings generated are conditional on also meeting quality benchmarks.
Comprehensive Primary Care Plus Initiative	Process, outcome, patient experience, efficiency	This demonstration gives providers a monthly care management fee and a performance-based incentive payment. Providers continue to receive payments under the Medicare Physician Fee Schedule, but those payments are reduced for Track 2 providers as the Centers for Medicare & Medicaid Services shifts to making payments as a quarterly lump sum.

Note: This table excludes quality reporting programs for Medicare Advantage and Medicare Part D plans because they are outside the scope of this paper.