CHAPTER 6

Report on comparing quality among Medicare Advantage plans and between Medicare Advantage and fee-for-service Medicare

RECOMMENDATIONS

6-1 The Secretary should define electronic health record "meaningful use" criteria such that all qualifying electronic health records can collect and report the data needed to compute a comprehensive set of process and outcome measures consistent with these recommendations. Qualifying electronic health records should have the capacity to include and report patient demographic data such as race, ethnicity, and language preference.

COMMISSIONER VOTES: YES 15 • NO 0 • NOT VOTING 0 • ABSENT 2

6-2 The Secretary should collect, calculate, and report quality measurement results in Medicare Advantage at the level of the geographic units the Commission has recommended for Medicare Advantage payments, and calculate fee-for-service quality results for purposes of comparing Medicare Advantage and fee-for-service using the same geographic units.

COMMISSIONER VOTES: YES 15 • NO 0 • NOT VOTING 0 • ABSENT 2

6-3 The Secretary should have all health plan types in Medicare Advantage report on the same basis, including reporting measures based on medical record review, and the Congress should remove the statutory exceptions for preferred provider organizations and private fee-for-service plans with respect to such reporting.

COMMISSIONER VOTES: YES 15 • NO 0 • NOT VOTING 0 • ABSENT 2

6-4 The Secretary should collect and report the same survey-based data that are collected in Medicare Advantage through the Health Outcomes Survey for the Medicare fee-for-service population, unless the Secretary determines that such data cannot meaningfully differentiate quality among Medicare Advantage plans and between fee-for-service and Medicare Advantage.

COMMISSIONER VOTES: YES 15 • NO 0 • NOT VOTING 0 • ABSENT 2

6-5 The Secretary should expeditiously publish specifications for forthcoming Medicare Advantage plan encounter data submissions to obtain the data needed to calculate patient outcome measures.

COMMISSIONER VOTES: YES 15 • NO 0 • NOT VOTING 0 • ABSENT 2

6-6 The Secretary should calculate fee-for-service results for Healthcare Effectiveness Data and Information Set administrative-only measures for those measures the Secretary determines can provide a valid comparison of the two sectors.

COMMISSIONER VOTES: YES 15 • NO 0 • NOT VOTING 0 • ABSENT 2

6-7 The Secretary should develop and report on additional quality measures for Medicare Advantage plan and Medicare Advantage–to–fee-for-service comparisons that address gaps in current quality measures.

COMMISSIONER VOTES: YES 15 • NO 0 • NOT VOTING 0 • ABSENT 2

6-8 The Congress should provide the Secretary with sufficient resources to implement the Commission's recommendations in this report.

COMMISSIONER VOTES: YES 15 • NO 0 • NOT VOTING 0 • ABSENT 2

CHAPTER 6

Report on comparing quality among Medicare Advantage plans and between Medicare Advantage and fee-for-service Medicare

Chapter summary

In recent years, the Commission has made a number of recommendations on quality reporting and quality-related payment adjustments in both the Medicare Advantage (MA) and traditional Medicare fee-for-service (FFS) programs. In response to a congressional mandate, this chapter contains additional recommendations on quality measurement and reporting in Medicare. Specifically, Section 168 of the Medicare Improvements for Patients and Providers Act of 2008 (MIPPA) requires the Commission to submit a report to the Congress by March 31, 2010, about measures for comparing quality and patient experience in the MA and FFS programs, with the goal of collecting and reporting such measures by the year 2011. MIPPA requires that the report:

- address methods for comparing quality among MA plans as well as between the MA and FFS programs,
- address issues in public reporting and benchmarking, and
- include recommendations for legislative or administrative changes as the Commission finds appropriate.

Any changes the Commission recommends in March 2010 would have to be implemented immediately for collection and reporting of measures in 2011. CMS, health plans, and other involved entities need as much lead time as possible to implement changes and be prepared for data collection

In this chapter

and reporting in that one-year time frame. Thus, we have taken an incremental approach, building on current measurement systems and data sources to improve quality comparisons in the short term—by 2011. For the longer term—that is, by 2013 and beyond—we recommend ways to expand current quality measurement and reporting systems where appropriate and to fill in gaps in the current measurement sets, including the use of outcome measures to compare MA and FFS in local geographic areas. We also recommend leveraging the capabilities and increased use of health information technology, which will be supported by Medicare payment incentives beginning in 2011, to facilitate improvements in quality measurement.

Medicare currently uses three systems to measure and compare quality across MA plans and track changes over time:

- Healthcare Effectiveness Data and Information Set (HEDIS®), which measures clinical processes and intermediate clinical outcomes;
- Consumer Assessment of Healthcare Providers and Systems (CAHPS®), which primarily measures patients' experiences of care delivered through their plans and providers; and
- Health Outcomes Survey (HOS), which measures changes in beneficiaries' self-reported physical and mental health status over time.

Are comparable data sources and measures available for the FFS program? The MA CAHPS survey has the most direct analogue in the FFS CAHPS survey, which CMS currently fields to a sample of beneficiaries in FFS Medicare. A limited number of HEDIS measures used in MA are also used to compute quality measures in FFS Medicare, and some of the HOS survey questions have been fielded as a component of the FFS CAHPS survey.

On the basis of our findings, the Commission makes eight recommendations. They address the use of electronic health records, the geographic unit of analysis for quality comparisons, uniformity in quality data reporting requirements, comprehensiveness of quality measures, and the issue of whether there are sufficient dedicated resources for CMS. Although the resources required to implement these recommendations are likely to be substantial, we believe it is important to beneficiaries, providers, and policymakers that comparisons on quality be as accurate and reliable as possible. The unintended consequences of incomplete or flawed comparisons would be detrimental to the goal of improving quality across Medicare. ■

MIPPA Section 168

SEC. 168. MEDPAC STUDY AND REPORT ON QUALITY MEASURES.

(a) STUDY.—The Medicare Payment Advisory Commission shall conduct a study on how comparable measures of quality and patient experience can be collected and reported by 2011 for the Medicare Advantage program under part C of title XVIII of the Social Security Act and the original Medicare fee-for-service program under parts A and B of such title. Such study shall address technical issues, such as data requirements, in addition to issues relating to appropriate quality benchmarks that— (1) compare the quality of care Medicare beneficiaries receive across Medicare Advantage plans; and (2) compare the quality of care Medicare beneficiaries receive under Medicare Advantage plans and under the original Medicare fee-for-service program.

(b) REPORT.—Not later than March 31, 2010, the Medicare Payment Advisory Commission shall submit to Congress a report containing the results of the study conducted under subsection (a), together with recommendations for such legislation and administrative action as the Medicare Payment Advisory Commission determines appropriate. ■

Introduction: Quality measurement and reporting

The Commission has long been interested in health care quality reporting and in creating links between Medicare provider payments and quality performance. We have made a number of recommendations on quality reporting and quality-related payment adjustments in both the Medicare Advantage (MA) and traditional Medicare fee-for-service (FFS) programs. Section 168 of the Medicare Improvements for Patients and Providers Act of 2008 (MIPPA) mandates that the Commission examine existing quality measures and make further recommendations.

The MIPPA mandate

Section 168 of MIPPA requires the Commission to submit a report to the Congress by March 31, 2010, about measures for comparing quality and patient experience in MA with the traditional FFS program, with the goal of collecting and reporting such measures by 2011 (see text box). The report should address ways to compare quality among MA plans as well as between MA and FFS, examine issues in reporting and benchmarking, and include recommendations for needed legislative or administrative changes.[1]

Previous Commission recommendations

The Commission has been concerned for many years that the Medicare program pays MA plans and providers in the FFS program without regard to the quality of the care that is provided. We have made a series of recommendations to allow the program to differentiate payments based on quality measures.

One potential use for the information on quality is in connection with pay-for-performance (P4P) systems that reward higher or improved quality. The Commission has recommended P4P for many of Medicare's payment systems, including MA, along with recommendations that would facilitate an MA-to-FFS comparison. Specifically, the Commission has recommended that:

- The Congress should establish a quality incentive payment policy for all MA plans and that CMS implement an incentive program to reward higher quality plans. Under this policy, CMS would create a reward pool from a small percentage of plan payments and redistribute it based on plans' performance attainment and improvement on quality indicators (Medicare Payment Advisory Commission 2004).
- CMS should require providers who perform laboratory tests to submit laboratory values, using common vocabulary standards (Medicare Payment Advisory Commission 2005a).
- The Secretary should calculate clinical quality measures for the FFS program that would permit CMS to compare FFS with MA (Medicare Payment Advisory Commission 2005b).

- The Congress should set the MA benchmark payment amounts that CMS uses to evaluate plan bids at 100 percent of FFS costs and redirect Medicare's share of savings from bids below the benchmarks to a fund that would redistribute the savings to MA plans on the basis of quality measures (Medicare Payment Advisory Commission 2005b).
- The Congress should change the current county-based MA payment areas to metropolitan statistical areas (as long as they do not cross state boundaries) and National Center for Health Statistics health service areas for a state's nonmetropolitan areas (Medicare Payment Advisory Commission 2005b).

In addition to these recommendations, the Commission noted in its June 2009 report that, during a transition period in which MA plan payments would fall as benchmarks were reduced from their current high levels (in relation to FFS), higher quality plans should receive higher payments than other plans. After the transition, if plans demonstrated that their quality was better than FFS, they could receive payment rates that were higher than FFS (Medicare Payment Advisory Commission 2009b).

In this chapter, we build on those recommendations and, consistent with the MIPPA mandate, focus on comparing quality of care among MA plans and between MA and FFS. We seek to accomplish three goals:

- enable CMS to better manage the Medicare program,
- provide a basis for differentiating payments based on quality, and
- provide beneficiaries with better information for making more informed choices among the MA plans and the FFS program.

Process and outcome measures for assessing quality

The metrics for assessing health care quality include process, intermediate outcome, and outcome (including patient experience) measures. In fulfilling the MIPPA mandate, we sought to identify those measures that can be clearly defined, practically collected, and meaningfully interpreted, taking into account the strengths and weaknesses of each kind of measure.

Process measures often focus on a single dimension of care for a specific condition—for example, the share of patients with a diagnosis of diabetes who received at least one blood glucose test in the measurement year. Process measures assess whether a specific test, treatment, or other intervention was delivered to patients for whom the process is indicated. Process measures can be relatively straightforward to define and measure in that information can be obtained from administrative data sets such as claims; both MA plans and CMS have experience using them.

Intermediate outcome measures indicate whether patients diagnosed with a particular chronic condition such as diabetes or hypertension are achieving improvement of a specific abnormal physiologic function (e.g., blood glucose or blood pressure) attributable to their condition. Intermediate outcome measures rely on actual test results to evaluate whether specific clinical treatment objectives were achieved, such as the percentage of diabetes patients who had low-density lipoprotein cholesterol levels below a specified target value during the measurement period. Although these indicators do not capture the end results of care, they can be important ongoing indicators of whether treatment for a specific condition is being prescribed by providers and adhered to by patients.

Outcome measures reflect the result of care, from either a clinical or a patient-centered perspective (Institute of Medicine 2006). Several types of outcomes are used as quality indicators, including mortality rates, hospital admission and readmission rates, and patient-centered measures, such as surveys of patients' experiences with the health care system or their self-perceived health status (Institute of Medicine 2006, National Quality Forum 2009). Outcome measures such as mortality and readmission rates provide an integrated assessment of quality because they reflect the result of multiple care processes provided by all health care providers involved in the patient's care. They also focus attention on much-needed system-level improvements, because achieving the best patient outcomes often requires carefully designed care processes, teamwork, and coordinated action on the part of many providers (National Quality Forum 2009). Patient experience measures are inherently subjective by nature, but they capture an important patient-centered dimension of quality not available elsewhere (Medicare Payment Advisory Commission 2004).

Near-term and long-term feasibility of quality measures

The MIPPA provision specifies a very short time between the March 2010 publication of this report and the initial

reporting of improved measures in 2011. From a practical point of view, any changes the Commission recommends would have to be implemented immediately to meet the 2011 time frame. CMS, health plans, and other involved entities will need lead time to implement data collection changes on services rendered in 2010 for measure reporting in 2011. To the extent that implementing a recommendation requires aggregating quality indicators based on currently collected data, those results could be reported in 2011. In contrast, for recommendations that involve a change in the data collection processes of MA plans during the course of the year (e.g., requiring the collection of new data and therefore incurring new costs that had not been anticipated in MA plan bids), data collection could begin in 2011 but reporting would not be possible until 2012.[2]

For the longer term—that is, by 2013 and beyond—we recommend ways to expand current quality measurement and reporting systems and to fill in gaps in the current measurement sets, including using MA plan encounter data and FFS claims data to calculate and compare outcome measures—such as hospital admission rates, readmission rates, and mortality rates—for MA plans and the local FFS Medicare beneficiary population. We also recommend leveraging the capabilities and increased use of health information technology (HIT), which will be supported by Medicare payment incentives beginning in 2011, to facilitate improvements in quality measurement in both MA and FFS Medicare.

We considered a variety of quality measures that would enable comparisons of quality among plans and between programs. We sort current measures by their use in MA and FFS and examine additional measures that would be useful but feasibly could be implemented only in the longer term. (A more detailed description of quality measurement systems used in MA is provided in the online appendix to this chapter at http://www.medpac.gov.)

Quality measurement systems used in Medicare Advantage

CMS uses three systems in MA to measure and compare quality across plans. The cost of each system is primarily borne by MA plans. The systems are:

- Healthcare Effectiveness Data and Information Set (HEDIS®)[3] —HEDIS measures are based on administrative data, such as claims and encounter data, supplemented with clinical data extracted from medical records for certain measures. HEDIS measures are either process or intermediate outcome measures. HEDIS is maintained by the private, not-for-profit National Committee for Quality Assurance (NCQA) and is used for Medicare, Medicaid, and many commercial health plans. NCQA works with CMS to adapt HEDIS measures to the Medicare population. MA plans have been reporting selected HEDIS measures since 1997. NCQA has also worked with CMS to take the plan-level HEDIS metrics and apply them at the physician practice level.

- Consumer Assessment of Healthcare Providers and Systems (CAHPS®)—CAHPS is a set of patient experience surveys administered to Medicare beneficiaries in both MA and FFS. CAHPS provides information on respondents' personal experiences interacting with their health plan and health care providers. CAHPS results are used to measure quality from the patient's perspective across six domains: quick access to care of any type, access to needed care without delays, effectiveness of physician communication, health plan information and customer service, overall rating of health care quality, and overall rating of health plan quality. CAHPS was developed by the Agency for Healthcare Research and Quality (AHRQ) and the MA version was first fielded in 1997. Components of the CAHPS survey are included in HEDIS reporting. CMS also fields a version of CAHPS in FFS Medicare. Many researchers have used CAHPS to compare quality between MA and Medicare FFS (e.g., Keenan et al. 2009, Landon et al. 2004).

- Health Outcomes Survey (HOS)—HOS is a longitudinal survey of self-reported health status among MA plan enrollees over a two-year period. For each MA plan, randomly selected enrollees are surveyed in a given year and then resurveyed two years later about perceived changes in their physical and mental health. The beneficiaries' physical and mental health status is categorized as better, the same, or worse than expected, based on a predictive model that takes into account risk adjustment factors to determine expected results. When results are reported, a plan is deemed to have better or poorer outcomes if the plan's results on the physical or mental health measures are significantly different from the national average across all plans. (Components of HOS results are included in HEDIS reporting.)

Quality measurement systems currently used in fee-for-service Medicare

CMS currently measures and compares quality in several fee-for-service payment systems, including inpatient hospitals and outpatient hospital departments, physicians and other eligible professionals, skilled nursing facilities, home health agencies, and dialysis facilities.

Inpatient hospitals. Hospitals that are paid under the inpatient prospective payment system (IPPS) have a financial incentive to participate in the Reporting Hospital Quality Data for Annual Payment Update (RHQDAPU) program. The Congress has authorized CMS to reduce the annual IPPS market basket update by 2 percentage points for any eligible hospital that does not successfully report the annually designated quality measures. In fiscal year 2007, nearly 95 percent of eligible hospitals participated successfully in the reporting program (Centers for Medicare & Medicaid Services 2008b).

The RHQDAPU program included 30 performance measures in fiscal year 2009 and includes 42 in fiscal year 2010. Almost all the process measures require medical record data abstraction by participating hospitals, including treatment of acute myocardial infarction (AMI), heart failure (HF), and pneumonia; surgical care improvement; patient safety; and nursing-sensitive care. Patient experience measures are based on data collected by hospitals through the Hospital Consumer Assessment of Healthcare Providers and Systems (H–CAHPS®) patient survey. These data include information about all patients served by the hospital, regardless of payer. For the RHQDAPU outcome measures (e.g., mortality and readmission rates for selected conditions), hospitals do not have to report data to CMS; instead, CMS calculates these measures with Medicare claims data. Outcome measures calculated by CMS include in-hospital and 30-day postdischarge mortality and complication rates for selected conditions and procedures, as well as 30-day readmission rates for patients with HF. Most hospital quality measures gathered through the RHQDAPU program are published on the Hospital Compare website (www.hospitalcompare.hhs.gov).

Hospital outpatient departments. In the Tax Relief and Health Care Act of 2006 (TRHCA), the Congress required CMS to establish a quality data reporting program for hospital outpatient care. Similar to the inpatient quality data reporting program, hospitals that fail to report on the designated outpatient quality measures incur a reduction of 2 percentage points in their annual outpatient prospective payment system payment rate update. In fiscal year 2009, the Hospital Outpatient Quality Data Reporting Program had 11 measures, including 5 measures related to care for patients with AMI, 2 related to preventing surgical infection, and 4 related to use of certain imaging procedures.

Physicians and other professionals. In TRHCA, the Congress authorized CMS to establish the Physician

(continued next page)

CMS recently began requiring MA special needs plans (SNPs) to report on a set of quality measures that apply only to plans of this type.[4] These "SNP-only measures" include:

- Five measures reported through HEDIS—advanced care planning, functional status assessment, medication review, pain screening, and medication reconciliation postdischarge (Centers for Medicare & Medicaid Services 2009a).[5]
- Six structural measures (i.e., measures of whether a plan has implemented policies and procedures to achieve specified goals)—complex case management, improving member satisfaction, clinical quality improvements, care transitions, relationship with member's nursing facility (institutional SNPs only), and coordination of Medicare and Medicaid benefits (dual eligible SNPs only) (National Committee for Quality Assurance 2009c).

Quality measurement systems currently used in fee-for-service Medicare

Quality Reporting Initiative (PQRI), which establishes a financial incentive for eligible professionals to participate in a voluntary quality reporting program. In 2009, eligible professionals who successfully met PQRI reporting requirements received a bonus payment equal to 2 percent of their total allowed charges for covered services payable under the Medicare physician fee schedule during the reporting period.

CMS does not publish PQRI results for individual physicians or physician groups but makes the results available to each physician or group. However, the Medicare Improvements for Patients and Providers Act of 2008 requires CMS to publish on the www.medicare.gov website the names of the physicians and group practices that satisfactorily submitted data on quality measures under PQRI.

Skilled nursing facilities and home health agencies. In 1999–2000, CMS required skilled nursing facilities (SNFs) and home health agencies (HHAs) to begin routinely collecting and submitting patient assessment data as a condition of participation in Medicare. The patient assessment instrument used to collect and report performance data by SNFs is the Minimum Data Set (MDS), and the corresponding instrument for HHAs is the Outcome and Assessment Information Set (OASIS). In 2002–2003, CMS used its existing statutory authority to publish on its website the SNF and HHA quality measures that CMS calculated from submitted MDS and OASIS data. These measures do not require SNFs or HHAs to submit any information beyond what they must submit through their respective patient assessment instruments. The SNF measures report on various quality indicators associated with common clinical conditions among SNF patients and how well SNFs help their patients regain or maintain their ability to function. Similarly, the HHA measures indicate how well HHAs help their patients regain or maintain their ability to function by using indicators of physical health status and how well people can perform activities of daily living, as well as utilization measures, such as hospital admissions and use of emergent care.

Dialysis facilities. CMS currently uses 22 measures to monitor the quality of care delivered to patients with end-stage renal disease (ESRD). The topic areas for these measures are anemia, dialysis adequacy, vascular access, mineral metabolism, influenza vaccination, mortality, and patient education, satisfaction, and quality of life. Currently, public reporting is limited to three measures—hematocrit level, urea reduction ratio, and mortality—that are available for 100 percent of the ESRD population. Data on other measures are collected from a 5 percent random sampling of the ESRD population. In October 2008, CMS implemented new conditions for coverage that all Medicare-participating dialysis providers must meet. The new conditions require that all dialysis facilities electronically submit their patients' clinical information to CMS via a web-based software application (CROWNWeb). According to CMS, CROWNWeb will allow the agency to publicly report more current quality data on the full set of ESRD quality measures. ■

Quality measurement in FFS Medicare

CMS currently uses a variety of quality measures to publicly report and track performance of the following types of FFS providers: inpatient hospitals, outpatient hospital departments, physicians and other eligible professionals, skilled nursing facilities, home health agencies, and dialysis facilities (Centers for Medicare & Medicaid Services 2009c) (see text box).[6] An important distinction between the quality measurement approach Medicare uses for MA and FFS is that quality in FFS Medicare is measured at the individual provider level, whereas quality in MA is measured at the plan level.

Additional quality measures considered

We examined additional measures that could be used to compare quality among MA plans and between MA and FFS. Some of these measures are beginning to be implemented by CMS and others will be more feasible in the future as new data sources become available, such as encounter data from MA plans and clinical data from electronic health records (EHRs). Some, but not all, of the measures described in this section have been endorsed by a multistakeholder consensus-based quality measurement entity such as the National Quality Forum. Nonetheless, the Commission intends that this report consider a wide

variety of approaches for improving quality measurement in MA and FFS Medicare, including measures that have strong research underpinnings and that will become more feasible to implement as clinical data become more readily captured and easily retrieved for quality measurement with the widespread use of EHRs.

Assessing Care of Vulnerable Elderly indicators The Assessing Care of Vulnerable Elders (ACOVE) project is a collaborative effort between RAND Health and Pfizer, Inc., to develop a set of quality indicators for the medical care provided to "vulnerable elders." This term was defined by the measure developers as community-dwelling individuals age 65 or older who have a relatively high near-term risk of death or functional decline (as assessed with a short standardized patient survey) and all patients aged 75 or older (Wenger et al. 2007). The most recent version of the measure set, ACOVE–3, contains 392 quality indicators covering 14 types of care processes and 4 domains of care: screening and prevention, diagnosis, treatment, and follow-up and continuity of care. Stakeholders with whom the Commission staff consulted and who were familiar with the ACOVE measures generally considered them superior measures for the target population. At the same time these experts thought the ACOVE measures were currently not feasible to implement on a wide scale, given their reliance on medical record data. However, most, if not all, of the information necessary to calculate the ACOVE measures could be efficiently extracted from EHRs if they were designed to capture and report the required data elements. We also note that the ACOVE measures are designed to apply specifically to patients age 65 or older (all of them have been validated for patients age 75 or older), and therefore they likely would not be appropriate for measuring the quality of care for Medicare beneficiaries under age 65.

Hospital readmission rates In its June 2007 report, the Commission discussed at length how hospital readmissions sometimes indicate poor care or missed opportunities to better coordinate care (Medicare Payment Advisory Commission 2007). CMS now uses Medicare claims data to calculate hospital-level 30-day risk-standardized readmission rates for 3 conditions: heart failure, acute myocardial infarction, and pneumonia. Hospitals do not need to report additional data for CMS to calculate these readmission rates. CMS began publishing hospital-level readmission rates for the three selected conditions on the Hospital Compare website in June 2009, and the agency plans to update them quarterly. Thirty-day readmission rates also are being tracked and used as a quality measure in the Medicare Acute Care Episode demonstration, in which hospitals and physicians in certain communities form a single accountable entity that accepts a bundled payment for designated orthopedic and cardiac procedures. CMS currently does not gather MA plan encounter data that would enable the calculation of readmission rates for MA plans, either at the aggregate sector level (i.e., across all plans) or at the individual plan or contract level.

Hospital admission rates for ambulatory care sensitive conditions The Agency for Healthcare Research and Quality (AHRQ) has developed a set of prevention quality indicators (PQIs) that are outcome measures designed to calculate rates of potentially preventable hospitalizations for specific ambulatory care sensitive conditions (ACSCs) in a given geographic area or population (such as enrollees in a health plan). ACSCs include conditions such as diabetes, congestive heart failure (CHF), and chronic obstructive pulmonary disease, for which high-quality outpatient care can prevent the need for hospitalization or for which early intervention can prevent complications or more severe disease (Agency for Healthcare Research and Quality 2007b). For those PQIs that are appropriate for the Medicare population, CMS could calculate them for Medicare FFS in a given geographic area with data from the Medicare Provider Analysis and Review (MedPAR) files that the agency compiles annually. CMS could also calculate PQIs for MA plans that submitted encounter data with the same data elements in the same level of detail as MedPAR files.[7]

Potentially preventable emergency department visits Researchers at the New York University Center for Health and Public Service Research have developed an algorithm for classifying emergency department (ED) visit data in four basic categories of use:

- nonemergent—cases in which immediate care is not required within 12 hours;
- emergent–primary care treatable—cases in which care is needed within 12 hours but could be provided in a typical primary care setting;
- emergent–ED care needed: preventable or avoidable—cases in which immediate care in an ED setting is needed for a condition that could have been prevented or avoided with timely and effective ambulatory care; and

- emergent–ED care needed: not preventable or avoidable—cases in which immediate care in an ED setting is needed for a condition that could not have been prevented or avoided with ambulatory care.

The algorithm also identifies visits with a primary diagnosis involving mental health, substance abuse, or injury (Billings 2003). The algorithm is designed to be used with administrative data sources, such as ED discharge data.[8]

Mortality rates AHRQ has developed and maintains a set of inpatient quality indicators (IQIs) that can be used with hospital inpatient discharge data (such as the information CMS currently collects for all FFS Medicare hospitalizations) to calculate mortality rates for certain conditions and medical procedures at the level of individual hospitals. These measures apply to inpatient conditions for which mortality rates have been shown to vary substantially across institutions and for which evidence associates high mortality with deficiencies in the quality of care (Agency for Healthcare Research and Quality 2007a). The current set of IQIs includes mortality rates for acute myocardial infarction, CHF, stroke, gastrointestinal hemorrhage, hip fracture, and pneumonia. These measures can be aggregated from the hospital level to geographic areas or health plans. To maximize the extent to which the IQIs reflect quality across a system of care (either among FFS providers in a geographic area or within an MA plan), Medicare could calculate mortality rates within 30 days of hospital discharge for patients with the specified conditions. As with the other outcome measures discussed in this section, the lack of MA plan encounter data means that CMS cannot calculate these measures for MA plans or the MA sector as a whole.

Burden of quality measurement and cost to CMS

Additional quality data collection and reporting represent an administrative burden on plans and providers that, while manageable, needs to be acknowledged. Plans already collect and report many quality measures. Some of the recommendations of this report involve minimal burden, consisting of a different manner of reporting or aggregating data that are already being collected. New sources of information, consisting of data collected primarily for purposes other than quality monitoring, can be the source of quality measurement with minimal additional burden to plans and providers. For example, CMS already intends to collect encounter data from MA plans, from which it could calculate outcome measures, such as readmission and mortality rates, in ways that are comparable to outcome measures calculated from FFS claims data.

A greater level of burden arises when a measure requires more depth (e.g., needing to review medical records) or more breadth (e.g., needing to survey more beneficiaries to obtain a statistically sufficient sample size). Medical record review is an expensive and labor-intensive process for paper-based medical records. The burden of medical record review on plans and providers increases if the number of geographic reporting units is expanded, if more measures are developed that require medical record review, and if more plan types are asked to report on measures requiring medical record review. The burden would extend to providers not under contract to a health plan if preferred provider organizations (PPOs) and private fee-for-service (PFFS) plans are asked to report on care rendered by non-network providers.

The Commission recognizes the additional burden incurred by plans, providers, and beneficiaries associated with the recommendations in this report. Nevertheless, the effort is needed to ensure comparability of measures across MA plans and to provide a more complete picture of the quality of care beneficiaries receive in MA and FFS. Similarly, we recognize that the efforts to improve and extend the collection of quality information will require substantially more administrative resources for CMS, but these efforts are critical to making appropriate comparisons across plans and programs. The unintended consequences of inaccurate quality comparisons would be detrimental to policymakers, providers, and beneficiaries. With each of the recommendations made in this chapter, we identify their cost implications and we support a policy of designated funding for those efforts that would incur additional costs. The Commission does not typically estimate the impact of its recommendations on Medicare's administrative costs, but in this case we believe it is important to indicate the directional impact on Medicare administrative costs for our recommendations to emphasize the importance of CMS having adequate resources to carry them out.

Recommendations

As part of the MIPPA mandate, the Congress directed the Commission to make recommendations for legislative

TABLE 6–1 **Roadmap of recommendations**

Year	MA-to-MA comparison	MA-to-FFS comparison
2009–2010	**6-1:** Define EHR "meaningful use" to include data collection and reporting needed for comprehensive set of process and outcome measures with robust risk adjustment	
2011*	**6-2:** Compare quality using same geographic unit as MedPAC-recommended MA payment areas • CAHPS® and HEDIS®	**6-2:** Compare quality using same geographic unit as MedPAC-recommended MA payment areas • CAHPS®
	6-3: All MA plan types collect data and report HEDIS® measures on same basis	
2013		**6-4:** Implement HOS for FFS, unless the Secretary determines it cannot meaningfully differentiate between FFS and MA
	6-5: Compute limited set of outcome measures based on MA plan encounter data	**6-5:** Compute limited set of outcome measures based on MA encounter data and FFS claims
	6-6: Continue to compute all HEDIS® measures	**6-6:** Compute administrative-only HEDIS® measures that can be validly compared
	6-7: Expand scope of measure sets to fill current gaps in populations and conditions	
Concurrent with CMS implementation	**6-8:** Provide resources to CMS sufficient for implementing recommendations	

Note: MA (Medicare Advantage), FFS (fee-for-service), EHR (electronic health record), CAHPS® (Consumer Assessment of Healthcare Providers and Systems), HEDIS® (Healthcare Effectiveness Data and Information Set), HOS (Health Outcomes Survey).
*To the extent that a recommendation involves aggregating quality indicators based on currently collected data, those results could be reported in 2011. For recommendations that involve a change in MA plans' data collection processes during the course of the year (e.g., requiring the collection of new data and therefore incurring new costs that had not been anticipated in MA plan bids), data collection could begin in 2011 but reporting would not be possible until 2012.

or administrative changes that the Commission finds appropriate to improve comparisons of quality and patient experience measures among MA plans and between the MA and the FFS programs. Table 6-1 provides a chronological roadmap for implementing our recommendations. Below, we describe the recommendations and explain their rationale and impact on beneficiaries and providers.

Recommendation 6-1: Ensure EHRs can be used to evaluate quality

The absence of clinically detailed quality measurement tools that are based on medical record information is a fundamental limitation on the scope—and for many providers, the validity and "actionability"—of existing quality measures. Ideally, quality measures should incorporate clinically relevant longitudinal information on patients' visits, diagnoses, procedures, medications, and laboratory results (Hayward 2008, Shahian et al. 2007). Large-scale efforts to extract these data from paper-based medical records on a routine basis are not practical, given their cost and time-consuming nature (Institute of Medicine 2006).

Today, EHRs hold promise to provide detailed clinical data for quality measurement and improved risk adjustment (Kmetik et al. 2007, National Quality Forum 2008). The adequacy of risk adjustment for quality measures based on administrative data remains a major concern for providers and health plans, because administrative data lack clinical detail and systematically underrepresent patient comorbidities and other factors related to baseline risk (Institute of Medicine 2006).

New Medicare incentives authorized by the American Recovery and Reinvestment Act of 2009 (ARRA) are expected to accelerate the adoption and use of EHRs by hospitals, physicians, and integrated delivery systems in the United States. Sections 4101 and 4102 of ARRA provide Medicare bonus payments to eligible professionals who are "meaningful users" of certified EHRs by calendar

years 2011–2014 and for hospitals that are meaningful users of certified EHRs by fiscal years 2011–2015. Starting in 2015, eligible professionals and hospitals that are not meaningful users of certified EHRs will receive reduced Medicare payments.[9] The Congressional Budget Office estimates that the incentive mechanisms in ARRA will boost EHR adoption rates to about 70 percent for hospitals and about 90 percent for physicians by 2019 (Congressional Budget Office 2009).

In August 2009, a Department of Health and Human Services (HHS) federal advisory committee (the HHS HIT Policy Committee) issued nonbinding recommendations for the meaningful use qualification criteria (Health Information Technology Policy Committee 2009), and the Commission submitted a comment letter on the committee's proposal, strongly supporting the use of HIT to improve the quality and reduce the cost of care for Medicare beneficiaries (Medicare Payment Advisory Commission 2009a). In January 2010, CMS set forth a proposed set of meaningful use criteria for the Medicare HIT subsidies in a notice of proposed rule making (Centers for Medicare & Medicaid Services 2010). The Commission intends to submit a comment letter to CMS stating its support for the proposed criteria, which largely follow the HIT Policy Committee's recommendations and are consistent with Recommendation 6-1. The final criteria defining meaningful use for at least the first two years of the Medicare subsidy program (2011–2012) are expected to be issued by CMS by mid-2010.

RECOMMENDATION 6-1

The Secretary should define electronic health record "meaningful use" criteria such that all qualifying electronic health records can collect and report the data needed to compute a comprehensive set of process and outcome measures consistent with these recommendations *[Recommendations 6-2 through 6-7]*. Qualifying electronic health records should have the capacity to include and report patient demographic data, such as race, ethnicity, and language preference.

RATIONALE 6-1

The definition of EHR meaningful use—in connection with forthcoming Medicare subsidies for providers' adoption and use of EHRs—will have a major impact on EHRs' capabilities to collect and report data needed for quality measurement, including improved risk adjustment of outcome measures. The forthcoming meaningful use criteria should require the technical capacity to capture and report the data elements needed to implement the Commission's other recommendations in this chapter pertaining to improving MA plan-to-plan and MA-to-FFS quality comparisons.

EHRs should contain accessible information on relevant patient demographic data, such as race, ethnicity, and language preference. From the provider or plan perspective, for example, it is useful to know whether a person requires translation services during an encounter. Some of the demographic information may be obtained from CMS's administrative records if the reliability of race and ethnicity data can be improved (Eicheldinger and Bonito 2008). In such cases, a person-level identifier, such as an encrypted personal identification number, would be needed to link the demographic data in Medicare administrative records to each beneficiary's EHR. Some relevant demographic information—such as finer distinctions in race and ethnicity categories and patient language preferences—is more feasibly collected by providers during patient encounters and may also be included in EHRs.

The vision of quality measurement underlying this recommendation is that, for most quality measures, the measurement and comparisons of quality at the MA plan and FFS area level will involve the aggregation of data reported by individual providers from the EHRs they maintain for their patients. A small number of MA plans also will be considered meaningful users of HIT for the purposes of the Medicare subsidies (if they meet the criteria specified by CMS) and they also will maintain EHRs for their patients. But most beneficiaries in MA and FFS Medicare likely will continue to be served by hospitals and physicians that do not participate in that subset of highly vertically integrated MA plans. The intent of this recommendation is to ensure that the necessary quality data elements can be captured, reported, and aggregated for as many FFS and MA enrollees as possible to allow comparisons of quality between the two sectors.

IMPLICATIONS 6-1

Spending

- No additional CMS costs would be incurred beyond baseline costs to implement EHR meaningful use criteria.

Beneficiary and provider

- EHRs meeting meaningful use criteria would offer providers and beneficiaries information on the full

scope of quality measures with reporting by race or ethnicity, gender, and age group.

- There would be no provider costs beyond baseline spending to acquire and use EHR systems that meet CMS meaningful use criteria.

Recommendation 6-2: Revise the geographic unit for reporting

For both the MA-to-MA comparison and the FFS-to-MA comparison, evaluations should be made for the same geographic area at a level that is meaningful for beneficiary decision making, for CMS's evaluation of the comparative quality of each plan and each sector, and for purposes of benchmarking (evaluating current quality and change over time). Currently, quality results for MA plans generally are reported on a contract-wide basis. MA contracts often cover a wide geographic area, sometimes an entire state—as in the case of plans in California, Florida, and Texas. Those large areas may include many diverse health care markets. Reporting at a smaller geographic level would provide a better picture of relative quality among MA plans and between MA and FFS—which is important for benchmarking purposes.

In its June 2009 report, the Commission recommended the use of MA payment areas consisting of metropolitan statistical areas (as long as they did not cross state boundaries) and National Center for Health Statistics health service areas for a state's nonmetropolitan areas (Medicare Payment Advisory Commission 2009b). These alternative payment areas would replace the current county-based payment areas and more closely approximate insurance markets. Increasing the size of payment areas, decreasing the size of quality reporting (contract) areas, and making payment and reporting areas coincide would have two benefits, even though an increase in the number of reporting units for quality measures would be required. First, program management would be improved by making it possible to differentiate plans on quality and then translate those differences to increased (or decreased) payments in a pay-for-performance system. Second, beneficiaries would have better quality information on the plans they could join in their area and would have the opportunity to make more informed choices. Changes that would allow reporting at the smaller geographic level can be in place by 2011.[10]

Because the new reporting areas would be smaller than the current contract areas in some cases, there would be an additional burden on plans (which finance much of the cost of the quality reporting activities). Costs are very likely to increase because survey sample sizes will have to be increased for statistically valid reporting. CMS would have the burden of computing results for each geographic area for the FFS sector. CMS would also need additional resources to boost sample sizes in its surveys of FFS beneficiaries. Having adequate survey sample sizes would also be an issue in FFS.

Smaller geographic areas may have fewer enrollees—sometimes too few to yield statistically valid results for purposes of public reporting. Similarly, there are plans with small numbers of enrollees or in which the enrollment may be large but dispersed over many health care markets (as is particularly true of PFFS plans and regional PPO plans). In certain circumstances, the Secretary would have to develop alternative ways to evaluate and report on quality within geographic areas—for example, by using three-year rolling averages or otherwise aggregating the information in a statistically valid manner that provides useful and reliable information about the performance of one plan relative to another in an area and with respect to FFS in the area. In some cases, public reporting on the performance of a given plan in a particular area may not be feasible, and the only possibility may be to report on the plan's overall performance across all its markets.

RECOMMENDATION 6-2

The Secretary should collect, calculate, and report quality measurement results in Medicare Advantage at the level of the geographic units the Commission has recommended for Medicare Advantage payments, and calculate fee-for-service quality results for purposes of comparing Medicare Advantage and fee-for-service using the same geographic units.

RATIONALE 6-2

The current collection and reporting of most quality measures in MA occur at the level of the MA contract. Some MA contracts cover very wide geographic areas. Plans in California that cover much of the state report one set of statewide HEDIS results, for example, even though parts of California have very different health care markets, with different provider and plan characteristics in each geographic area.

To inform beneficiaries about the relative quality of MA plans and MA relative to FFS, comparisons should pertain to the geographic area where beneficiaries are making

choices. Using a smaller geographic area that is more consistent with the patterns of health care delivery would also facilitate CMS's quality monitoring and evaluation role in both MA and FFS.

IMPLICATIONS 6-2

Spending

- Substantial CMS administrative resources would be required.

Beneficiary and provider

- Beneficiaries' ability to compare plans and systems would be improved, but more beneficiaries would be included in surveys.
- Many plans would face additional costs because of an increase in the number of reporting units.

Recommendation 6-3: Level the playing field among MA plan types for HEDIS reporting

This recommendation pertains to the comparison among MA plans across different plan types (HMOs versus PPOs and PFFS plans). HEDIS reporting requirements for MA plans consist of process measures and intermediate outcome measures that are based on administrative data (claims data, encounter data, laboratory results, and EHRs), supplemented in some cases by information obtained from individuals' medical records. The latter type—the so-called "hybrid" measures—can include information drawn from a sample of plan enrollees' medical records as well as administrative data.

In the past, we expressed concern about the lack of a level playing field among MA plans in HEDIS reporting—because PFFS plans were not required to participate in HEDIS reporting and because not all plans report on the same basis (Medicare Payment Advisory Commission 2008). As of 2010, all MA plan types—HMOs, PPOs, and PFFS—have HEDIS requirements, but the requirements vary by plan type. PPOs and PFFS plans are exempted from reporting measures that are based exclusively on medical record review. Only HMOs and PPOs (not PFFS plans) are permitted to include data from medical records when they report on HEDIS hybrid measures that are based on a combination of administrative data and medical record review.[11]

Another difference among plan types in HEDIS reporting is that a statutory provision limits the reporting of PPOs to services rendered by network providers. A MIPPA provision added a similar limitation on PFFS and medical savings account plan reporting as of 2011—such plans have to report only on care provided through contracted providers (Centers for Medicare & Medicaid Services 2009b).[12]

To have uniform, comparable reporting across MA plans, which is a prerequisite for benchmarking plan performance, reporting standards and practices need to be the same across plans. All plan types should report results for all providers, and all plans should use medical record review as appropriate to report results. These changes are feasible for data collection occurring in 2011 for reporting in 2012.

RECOMMENDATION 6-3

The Secretary should have all health plan types in Medicare Advantage report on the same basis, including reporting measures based on medical record review, and the Congress should remove the statutory exceptions for preferred provider organizations and private fee-for-service plans with respect to such reporting.

RATIONALE 6-3

Requiring all plans to report using the same methodology enables a valid plan-to-plan comparison across all HEDIS measures, including intermediate outcome measures, which involve medical record review. A plan should report on all services its enrollees receive—regardless of whether providers are under contract.

IMPLICATIONS 6-3

Spending

- CMS would incur costs in processing more data than would otherwise be reported.

Beneficiary and provider

- Beneficiaries would have better information to compare MA plans.
- Plans would incur additional costs in reporting on measures requiring medical record review. Some providers could incur additional costs in providing medical records for review.

Recommendation 6-4: Enhance the Health Outcomes Survey for MA and FFS

The HOS is a longitudinal survey of self-reported health status among MA plan enrollees over a two-year period.

There currently is no HOS in FFS Medicare. The survey's methodology could be improved to make it a better tool for comparing MA plans and for eventual use in FFS Medicare to compare outcomes between MA and FFS. The problem is that HOS produces results that often show no significant difference among most plans in enrollee outcomes. (HOS results are reported differently at the www.medicare.gov website, with more differentiation among plans, including in the CMS star rating system for plan quality.[13] This issue is discussed in the online appendix to this chapter.)

NCQA is working with CMS to study why the current HOS methodology identifies only a few outlier plans, with a view toward recommending potential changes to the methodology and developing new methods and processes. NCQA and CMS will examine the current case mix variables, current statistical methods used in HOS, and current criteria for establishing outliers. NCQA has noted that one issue is that the need for a two-year change score, which is the basis of judging outcomes, limits the number of enrollees with reportable HOS results and "may contribute to the lack of variation and usefulness of the measure" (National Committee for Quality Assurance 2009a). Under the current methodology, if CMS were to field the equivalent of the HOS in the FFS sector, it is uncertain whether the results would show statistically valid differences between MA and FFS or across FFS.

Implementation of the HOS in FFS would be a major undertaking that would involve a lengthy planning and start-up period. For this reason, and because the HOS involves an initial survey and a follow-up survey two years later, implementation of such a survey in FFS would not produce results until well after 2011. Changes to the methodology for making comparisons among MA plans could be implemented by 2011.

RECOMMENDATION 6-4

The Secretary should collect and report the same survey-based data that are collected in Medicare Advantage through the Health Outcomes Survey for the Medicare fee-for-service population, unless the Secretary determines that such data cannot meaningfully differentiate quality among Medicare Advantage plans and between fee-for-service and Medicare Advantage.

RATIONALE 6-4

The HOS could be a valuable tool in program management, quality improvement, and beneficiary education. Work should start on a FFS-to-MA comparison, but the Secretary should investigate whether greater distinctions can be drawn among MA plans and whether meaningful differences can be reported between MA and FFS. Currently, HOS results in MA do not show clear distinctions among plans. Extensive resources would be required to conduct the HOS across the FFS sector.

IMPLICATIONS 6-4

Spending

- Substantial CMS administrative resources would be required if the HOS is expanded to FFS beneficiaries.

Beneficiary and provider

- Beneficiaries' ability to compare plans and systems would be improved, but more beneficiaries would be included in surveys, increasing the response burden.
- There would be no implications for plans and providers.

Recommendation 6-5: Use MA plan encounter data to evaluate quality

Medicare needs encounter data from MA plans so that it can use outcome measures to assess and compare the quality of inpatient and ambulatory care in MA and FFS Medicare. Patient encounter data are collected by health plans from the health care facilities and professionals who provide services to the plan's members. Encounter data may be derived from claims submitted by providers to the plan (including "zero-pay" or "no-pay" claims, which are used not to pay a provider but only to generate an encounter record), or the necessary data elements may be extracted from patient-level EHR systems maintained by providers.

As discussed on page 316, we examined four types of outcome measures that could be used as quality indicators for MA plans and for FFS Medicare within a designated geographic area:

- hospital readmission rates for conditions in which clinical evidence suggests that appropriate discharge planning and postdischarge follow-up can prevent readmission;
- hospital admission rates for ACSCs;
- potentially preventable ED visits; and
- mortality rates during or within up to 30 days after a hospital stay for patients diagnosed with specific

conditions, such as a heart attack, heart failure, or pneumonia.

These measures could be computed for FFS today using existing claims and hospital discharge record data. Medicare currently cannot use these measures to assess and compare quality among MA plans and between MA and FFS Medicare because the necessary encounter data for MA enrollees are not available.

RECOMMENDATION 6-5

The Secretary should expeditiously publish specifications for forthcoming Medicare Advantage plan encounter data submissions to obtain the data needed to calculate patient outcome measures.

RATIONALE 6-5

Outcome measures are important indicators of the quality of care provided to MA plan members and FFS beneficiaries in a given geographic area. Four types of outcome measures can be calculated for FFS Medicare with available claims data and could be calculated for MA plans if plans were required to submit the necessary encounter data. CMS intends to require MA plans to submit encounter data starting in 2012, which presents an opportunity to request the encounter data elements needed to compute the specified outcome measures for MA plans, enabling comparisons between MA plans and between MA and FFS Medicare by 2013.

IMPLICATIONS 6-5

Spending

- Little or no additional administrative costs would be incurred above the costs already assumed in the agency's budget for collecting encounter data from MA plans beginning in 2012.

Beneficiary and provider

- For beneficiaries, important information on patient outcomes would be available when comparing MA plans and comparing MA with FFS in their local area.
- Providers and plans could incur costs above those assumed for the planned 2011 encounter data collection and reporting.

Recommendation 6-6: Compute selected HEDIS measures for FFS Medicare

Some measures in the HEDIS MA data set could be the basis of an MA-to-FFS comparison if HEDIS measure specifications were used to compute FFS results derived from claims data. This approach has been used by CMS on a pilot basis in its Generating Medicare Physician Quality Measurement Results project and by the Dartmouth Atlas group for the Robert Wood Johnson Foundation's Aligning Forces for Quality project (Fisher et al. 2008). However, we have a number of concerns about whether FFS and MA HEDIS measures could be truly comparable without some adjustments to the measure specifications and to the populations being compared.

The first concern is about the use of hybrid HEDIS measures, which are those that include the use of administrative data and medical record review in MA. FFS claims data alone are insufficient to compute a measure comparable to the MA result. Even if the MA results for hybrid measures were limited to an administratively determined rate (i.e., without medical record review), the administrative rate from an MA plan is based on claims data, encounter data, pharmacy information, and in some cases electronic medical records—a richer source of information than FFS claims (even if they were combined with Medicare Part D pharmacy data). In light of expected new sources of information on quality indicators to compare FFS and MA (encounter data and EHRs) that would provide an equivalent type of information, it would be unreasonable to undertake a major effort to obtain what would end up being duplicative information from FFS through medical chart review for purposes of comparing MA results on hybrid HEDIS measures.

A second concern is that, even for the HEDIS derived only from administrative data, there can be material differences, unrelated to the quality of care, between a HEDIS rate reported by an MA plan and a FFS rate computed from claims data. In addition to MA plans' richer sources of administrative data, other factors would affect an MA-to-FFS comparison—namely, differences in populations and cost-sharing requirements that can affect utilization rates. The HEDIS measure for breast cancer screening is illustrative. The share of beneficiaries under age 65 is smaller in MA than in FFS. Because mammography screening rates are lower in the under-65 population, this factor would need to be taken into account for a valid comparison. The text box on p. 324 elaborates on this difference and the influence of cost sharing on the use of mammography.[14]

A third confounding factor in examining quality differences between MA and FFS is the potential for a spillover effect—that is, the effect an area's MA plans

How population distribution and cost sharing can affect Healthcare Effectiveness Data and Information Set measure comparability between Medicare Advantage and fee-for-service Medicare

Breast cancer screening rates were among the fee-for-service (FFS) results that CMS reported through its Generating Medicare Physician Quality Measurement Results (GEM) project, which was a CMS initiative in 2007–2008 that computed Healthcare Effectiveness Data and Information Set (HEDIS®) measures in FFS by geographic area using claims data (Centers for Medicare & Medicaid Services 2008a).

The average Medicare Advantage (MA) plan screening rate was about 10 percentage points higher than the FFS national average rate. For FFS, the GEM project reported the total rate across the 40- to 69-year age group for 2006–2007. MA plans reported total rates for the same age group as well as separate rates for the 40- to 52-year and 53- to 69-year age groups. The screening rates for the younger age group in MA plans were much lower than for the older group. If the same relationship held in FFS (lower rates for younger than for older women), the total rate reported for FFS in the GEM data would be understated in relation to the MA rate, because MA plans enroll a much smaller proportion of Medicare beneficiaries under age 65 than are enrolled in FFS Medicare. A more comparable measure for breast cancer screening would focus on the screening rates for women only in the age 65 or older group.

The mammography measure could demonstrate "value added" by MA plans relative to FFS. MA plans have the potential advantage of having greater systematic control over screening rates through telephone and mail reminders to beneficiaries plan wide, and plans can facilitate access to care (including, among some plans, providing transportation). In contrast, FFS Medicare tools for improving aggregate screening rates are more diffuse, relying on efforts such as public health campaigns and notifications by individual providers.

Another issue affecting data comparability involves cost-sharing differences between MA and FFS and among MA plans. Trivedi and colleagues examined cost sharing for mammograms in MA plans and found that "relatively small copayments were associated with significantly lower mammography rates among women who should undergo screening mammography according to accepted clinical guidelines" (Trivedi et al. 2008). Differences in screening rates that reflect cost-sharing differences also arise in FFS Medicare. Results in FFS can differ by geographic area when a large percentage of FFS Medicare beneficiaries have supplemental insurance—such as medigap, employer-based retiree health benefits, or Medicaid coverage—that reduces or eliminates FFS cost sharing for the services being measured. Benefit design and the richness of an individual's Medicare benefit package in MA or as supplemented in FFS can therefore have an effect on a HEDIS measure that is intended to show a difference in the quality of care that providers and health plans render in each program. ■

may have on FFS quality. The hypothesis is that, because many of the same providers treat patients covered under MA and FFS, any MA plan-driven quality improvements translate into changes in providers' practice patterns for patients treated in FFS as well, making it difficult to isolate the effect of quality improvements in one program or the other. The text box briefly discusses the spillover hypothesis. One benefit of CMS collecting comparable quality data on both the FFS and MA programs would be the opportunity to further test the validity of the spillover hypothesis.

Given these differential factors, some HEDIS-like measures may need adjustments to produce valid comparisons between MA and FFS.

RECOMMENDATION 6-6

The Secretary should calculate fee-for-service results for Healthcare Effectiveness Data and Information Set administrative-only measures for those measures the Secretary determines can provide a valid comparison of the two sectors.

Quality of care: The spillover effect

The literature on the potential for spillover between Medicare Advantage (MA) and fee-for-service (FFS) Medicare to affect quality is mixed. A number of researchers have found such an effect (as shown in multiple articles identified by Federman and Siu (2004)). Heidenreich and colleagues found that in areas with high HMO penetration (commercial and other enrollment), Medicare FFS beneficiaries were more likely to receive appropriate treatment with beta-blockers and aspirin following a heart attack—indicating a positive spillover effect (Heidenreich et al. 2002). Basu and Mobley, however, found that a county's managed care penetration (in commercial and Medicare HMOs) did not have a significant effect on preventable hospital admissions in any of the four states they examined (California, New York, Florida, and Pennsylvania) (Basu and Mobley 2007). Additional research is under way on the effect of MA spillover on quality (Harvard Medical School 2009).

Some activities that plans undertake (such as advising providers to contact enrollees to obtain tests and monitoring) would not have a direct spillover effect in FFS. Quality improvement activities in FFS may also "spill over" to benefit MA plans. For example, the efforts of FFS Medicare to evaluate the quality of providers (e.g., through Hospital Compare) can lead to improvements in provider quality across all sectors. ■

RATIONALE 6-6

HEDIS-like measures for FFS can be calculated in a straightforward manner with Medicare FFS claims data (including prescription drug event data from Part D) for those HEDIS measures that do not rely on medical record review and that the Secretary finds can yield valid comparisons of quality between the MA and FFS programs. CMS has computed such measures in the past and reported results at the ZIP code level. However, comparisons need to be viewed with caution because there are important differences between MA and FFS that affect the results. MA administrative data can include additional information not currently available in FFS administrative data systems. Differences between the populations and benefit design of the two programs should also be taken into account. The Secretary should ensure that the HEDIS-like measures in FFS that are compared with MA results reflect differences in quality and not other factors.

IMPLICATIONS 6-6

Spending

- CMS would incur administrative costs in computing and reporting the selected HEDIS measures for FFS Medicare.

Beneficiary and provider

- Beneficiaries' ability to compare plans and programs would improve.
- Providers and plans, which currently submit the data that would be used for these computations, would incur no additional costs.

Recommendation 6-7: Add new quality measures

An issue of concern with the current HEDIS measures is whether they are sufficiently comprehensive for Medicare beneficiaries. Among the set of 46 HEDIS measures for Medicare, 19 are drug related, but few non-drug-related measures apply to the oldest Medicare beneficiaries. For example, of the 6 intermediate outcome measures, only 1 applies to beneficiaries between 75 and 85 years of age, and none applies to people over 85.

Quality measures for diabetes provide a case study. The 9 HEDIS diabetes measures are reported only through age 75. However, Medicare Current Beneficiary Survey data indicate that about 20 percent of community-dwelling Medicare beneficiaries age 75 to 84 have diabetes (declining to 13.5 percent in the 85+ age category) (Adler 2008)). According to one estimate, nearly half the elderly with diabetes (44 percent) are not included in HEDIS diabetes measures (McBean et al. 2003). The HEDIS measure's cut-off at age 75 exists because beneficiaries in the older age groups require tailored, person-specific plans of care to deal with diabetes, precluding the use of uniform measures for these individuals (National Committee for

Quality Assurance 2009b). This problem exists for many conditions, and it is unclear how to overcome the problem for a population with so many comorbidities.

NCQA has been adding more measures for the very aged to the HEDIS data set. A subset of MA plans—the SNPs that serve the chronically ill, beneficiaries eligible for both Medicare and Medicaid (dual eligibles), and institutionalized beneficiaries—report an additional set of measures, which could apply to all MA plans. These measures can provide an indication of the value added that a plan can offer beyond the quality of care rendered by a plan's individual network providers.

Compared with a comprehensive set of process measures for geriatric care like the ACOVE indicators, HEDIS has few measures of quality for conditions prevalent among the Medicare population, such as treatment for chronic pain, dementia, end-of-life care, and malnutrition. Even when HEDIS includes clinical measures, the results may be of limited usefulness. For example, CMS excludes seven measures from the star rating system of overall plan quality because the incidence of the services being measured is too low to be statistically valid. All five HEDIS Medicare mental health measures (two for follow-up after inpatient mental health care and three for antidepression medication management) are excluded from the star system for this reason.

RECOMMENDATION 6-7

The Secretary should develop and report on additional quality measures for Medicare Advantage plan and Medicare Advantage–to–fee-for-service comparisons that address gaps in current quality measures.

RATIONALE 6-7

Expanding HEDIS's quality measures to cover a wider range of Medicare beneficiaries and more medical conditions would make the quality reports generated from HEDIS meaningful and actionable by plans and providers to improve the quality of care for beneficiaries with the specified characteristics, such as beneficiaries over age 75 and beneficiaries with disabilities. The addition of measures that assess plan functions, such as care coordination and medication management, would provide information on the value of quality improvements that plans offer in addition to the care rendered by a plan's network of individual providers.

IMPLICATIONS 6-7

Spending

- Additional administrative resources for CMS would be required.

Beneficiary and provider

- Beneficiaries with certain characteristics—such as older beneficiaries, those with disabilities, or those with certain chronic health conditions—would have access to quality information that is more pertinent to their health care needs.
- Providers and plans would incur cost increases for collecting and reporting data needed to compute new HEDIS measures.

Recommendation 6-8: Provide resources to CMS sufficient to implement other recommendations

The Commission is aware that implementation of the foregoing recommendations would require significant CMS administrative resources. Because of the analytic and labor-intensive nature of the tasks involved, this level of resources is needed to ensure that new quality measures developed and existing measures refined will produce accurate and reliable comparisons. Faulty comparisons would be detrimental to:

- the goals of policymakers who seek to pay MA plans and FFS providers differentially based on their relative performance on quality measures;
- plans and providers that seek to use Medicare quality reports for internal quality improvement efforts; and
- beneficiaries who need a reliable, objective source of information for comparing quality among plans and between FFS and MA.

RECOMMENDATION 6-8

The Congress should provide the Secretary with sufficient resources to implement the Commission's recommendations in this report.

RATIONALE 6-8

The resources required to implement Recommendations 6-1 through 6-7 are likely to be substantial. It is important to beneficiaries, plans, providers, and policymakers that quality comparisons between MA and FFS Medicare and among MA plans are accurate, as the unintended

consequences of faulty quality comparisons would be detrimental to Medicare beneficiaries, plans, and providers. It is unlikely that CMS would be able to implement the recommendations in this report with the necessary level of precision without additional administrative resources. For this reason, we believe dedicated resources are necessary. The Secretary should submit a budget proposal to the Congress that estimates the funding needed to implement the recommendations in this report.

IMPLICATIONS 6-8

Spending

- Additional costs would be incurred by taxpayers, beneficiaries, plans, or some combination of the three, depending on the funding approach selected by the Congress.

Beneficiary and provider

- Beneficiaries, plans, providers, and policymakers would have an improved ability to compare the quality of care among MA plans and between MA and FFS Medicare across several dimensions (process, outcome, and patient experience measures).

Conclusion: A set of recommendations to improve quality comparisons

The Commission recognizes that quality measurement and reporting must serve the needs of four distinct audiences: Medicare policymakers, health plans, health care providers, and Medicare beneficiaries. We emphasize that the recommendations presented in this report should be considered as a cohesive and interdependent set of actions that, if implemented in their entirety, will significantly improve policymakers' and beneficiaries' ability to compare the quality of care among MA plans and between MA and FFS Medicare. Medicare will be able to benchmark the performance of MA plans and FFS Medicare across multiple domains of quality—clinical processes, outcomes, and patient experience—to obtain a more complete picture of quality within appropriately sized geographic areas and to track changes over time. Health plans and providers will have more comprehensive and actionable information about the quality of the care they administer or deliver. More comprehensive quality measurement also should improve the public reporting of quality measures and enable beneficiaries to make more informed decisions. In future work, the Commission plans to explore in detail how Medicare beneficiaries use information about quality and other factors to make health care decisions, such as whether to enroll in an MA plan or which FFS providers to select in their community.

Our recommendations reflect the practical reality that CMS, health plans, and health care providers need as much lead time as possible to implement any changes to Medicare's current quality measurement and reporting methods. Therefore, we took the approach of adapting current measurement systems and data sources to start improving quality comparisons by 2011. By 2013, we recommend using a limited set of clinical process and outcome measures to compare MA and FFS, while working to increase the scope of quality measures available, ultimately leading to a more comprehensive, meaningful, and actionable set of measures. For the longer term, Medicare should take advantage of the coming increase in the adoption of HIT to improve the clinical relevance and robustness of the measures Medicare uses. Lastly, it is essential that CMS be provided with sufficient dedicated administrative resources to implement the package of recommendations in this report. ■

Endnotes

1 Benchmarking includes: evaluating performance in relation to a norm or expected level of performance and in relation to peers or similar entities, establishing an expected level of performance and tracking performance over time, determining the degree of improvement expected over time, and using data to distinguish among entities for purposes such as rewarding higher quality performance or correcting or sanctioning poorer performance. Benchmarking also includes a public reporting component in determining how to convey differences—for example, in the methodology that the National Committee for Quality Assurance uses in its national ranking of health plans or in the star ranking system that CMS uses for the Health Plan Compare website.

2 For example, if more enrollees need to be included in a beneficiary survey paid for by MA plans, CMS would convey information about the new or additional requirements in 2010 for implementation during the 2011 contract year, and the results would be reported in 2012.

3 HEDIS is a registered trademark of the National Committee for Quality Assurance. CAHPS is a registered trademark of the Agency for Healthcare Research and Quality.

4 SNPs are MA plans that can limit their enrollment to certain categories of beneficiaries. The three types of SNPs are those for dual eligibles (Medicare beneficiaries with Medicaid coverage), for beneficiaries residing in nursing facilities (institutional SNPs), and for beneficiaries with specific medical conditions.

5 SNPs also report on 12 standard HEDIS effectiveness-of-care measures if the SNP benefit package is a component of a larger MA contract. All applicable HEDIS measures are reported for all enrollees across the entire contract, but the 12 measures must be reported for each SNP benefit package within the contract. Some MA contracts consist only of SNP plans, in which case the MA plan reports all the HEDIS measures that any other plan would report.

6 CMS currently does not track quality measures for the following FFS provider types: ambulatory surgical center (ASC), independent rehabilitation facility, long-term care hospital, hospice, clinical laboratory, and durable medical equipment. In some of these cases (e.g., ASC and hospice), CMS is actively developing quality measures and would need to use a regulatory notice and comment process before implementing them.

7 CMS could calculate PQIs for MA enrollees in a very limited way today using the AHRQ Healthcare Cost and Utilization Project (HCUP) State Inpatient Databases, which identify Medicare hospital discharges as being for FFS Medicare or MA patients in 15 states. However, even for those 15 states where the HCUP databases distinguish between MA and FFS patients, the HCUP data do not identify the specific MA plans in which beneficiaries were enrolled when they were hospitalized, so plan-level measurement and comparisons with local FFS outcomes are not possible. There also appear to be other limitations in some of the HCUP databases that would prevent stratifying outcome measure results for specific groups of beneficiaries, such as by race or ethnicity.

8 The development work on the ED use classification algorithm was supported by the Commonwealth Fund, the Robert Wood Johnson Foundation, and the United Hospital Fund of New York (New York University Center for Health and Public Service Research 2009).

9 Eligible physicians who are not meaningful users of certified HIT systems by 2015 will see their Medicare payments reduced by the following amounts: 1 percent in 2015, 2 percent in 2016, and 3 percent in 2017 and each subsequent year. (The reductions are not cumulative; they are reductions of the amount the provider otherwise would have received in that year.) For 2018 and each subsequent year, if the proportion of eligible physicians who are meaningful EHR users is less than 75 percent, the payment reduction will further decrease by 1 percentage point from the applicable amount in the previous year, though the reduction cannot exceed 5 percent. The Secretary may, on a case-by-case basis, exempt eligible physicians (e.g., rural physicians who lack sufficient Internet access) from the payment reduction if it is determined that being a meaningful EHR user would result in significant hardship. Such exemptions may not be granted for more than five years (Congressional Research Service 2009).

10 Even if the metropolitan statistical area becomes the reporting unit for plans operating in urban areas, further refinements could be made to the reporting unit, as multiple benefit packages can be offered under one MA contract in the same metropolitan statistical area. For example, one MA organization could offer three HMO packages: one that has a very rich benefit package offered to employer group-sponsored retirees (subsidized by the former employer), a package for individuals with a high premium but minimal cost sharing, and a low-premium plan with high cost sharing. The benefit packages could also vary in the drug coverage offered, which can affect the ability of enrollees to adhere to a drug regimen—in turn affecting quality measurement results. Thus, a further refinement to quality reporting is to consider reporting at the level of the plan benefit package (as is done for SNPs for certain HEDIS measures). An analogue to such reporting in FFS would be to report results based on

different beneficiary characteristics, such as those with and without supplemental coverage through medigap, Medicaid, or employer-sponsored supplemental coverage.

11 As of 2010, CMS changed its past policy of not allowing PPOs to use medical record data in reporting HEDIS results (Centers for Medicare & Medicaid Services 2009a).

12 Some quality measures (e.g., in HEDIS) can be based on pharmacy data. For such measures, there is an unlevel playing field within MA. Most enrollees of MA plans obtain their Medicare Part D coverage through the MA plan, but PFFS plans are not required to offer drug coverage. PFFS enrollees in those circumstances can obtain Part D coverage from stand-alone prescription drug plans. To determine quality measures based on pharmacy data for such enrollees, data could be obtained from the prescription drug plans. However, not all beneficiaries in Medicare elect drug coverage, including those who have retiree drug coverage subsidized by Medicare and those who do not enroll in Part D at all. For these beneficiaries, drug-based quality measures are not available.

13 CMS has developed a measurement system of plan ratings in particular domains of quality, with plans awarded from one to five stars, in half-star increments, based on their performance in each domain, along with an overall rating for plan quality based on those domains. The domains include measures or results from HEDIS, CAHPS, HOS, appeals information from the independent review entity, plan disenrollment rates, and CMS's tracking of complaints and plan compliance activity (such as corrective action plans). With regard to HEDIS, CMS has removed from the star rating system several HEDIS measures owing to small numbers and the consequent lack of reliability of the measures. These measures are management of depression medication, mental illness measures, and persistence of beta-blockers after a heart attack. The star ratings and the source of the data are posted on CMS's website for public reporting: www.medicare.gov (the Health Options Compare site).

14 Another administrative-only HEDIS measure that would allow for a seemingly straightforward comparison between MA and FFS is the glaucoma screening measure. For MA, the HEDIS measure is the percent of Medicare enrollees age 65 or older, without a diagnosis of glaucoma, who were screened for glaucoma over the course of the year. In FFS Medicare, glaucoma screening is a covered benefit for high-risk beneficiaries (composed of individuals with diabetes, those with a family history of glaucoma, African Americans over the age of 50, and Hispanics age 65 or older). To have a valid comparison between the two sectors, the MA results would have to be adjusted to include only the high-risk Medicare FFS categories in the denominator.

References

Adler, G. 2008. Diabetes in the Medicare aged population, 2004. *Health Care Financing Review* 29, no. 3 (Spring): 69–79.

Agency for Healthcare Research and Quality, Department of Health and Human Services. 2007a. *Guide to inpatient quality indicators: Quality of care in hospitals—volume, mortality, and utilization, version 3.1*. Rockville, MD: AHRQ. March.

Agency for Healthcare Research and Quality, Department of Health and Human Services. 2007b. *Guide to prevention quality indicators: Hospital admission for ambulatory care sensitive conditions, version 3.1.* Rockville, MD: AHRQ. March.

Basu, J., and L. Mobley. 2007. Do HMOs reduce preventable hospitalizations for Medicare beneficiaries? *Medical Care Research and Review* 64, no. 5 (October): 544–567.

Billings, J. 2003. *Using administrative data to monitor access, identify disparities, and assess quality of the safety net. Tools for monitoring the health care safety net*. Rockville, MD: AHRQ. September.

Centers for Medicare & Medicaid Services, Department of Health and Human Services. 2008a. GEM documentation. http://www.cms.hhs.gov/GEM/Downloads/GEMImportantNotes.pdf.

Centers for Medicare & Medicaid Services, Department of Health and Human Services. 2008b. Reporting hospital quality data for annual payment update fact sheet. http://www.cms.hhs.gov/HospitalQualityInits/downloads/HospitalRHQDAPU200808.pdf.

Centers for Medicare & Medicaid Services, Department of Health and Human Services. 2009a. 2010 HEDIS, HOS and CAHPS measures for reporting by Medicare Advantage organizations. Note to Medicare Advantage quality contacts and Medicare compliance officers. December 2.

Centers for Medicare & Medicaid Services, Department of Health and Human Services. 2009b. Issuance of the 2010 Call Letter. Note to all Medicare Advantage HMOs, PPOs, PFFS, §1876 cost contractors, and special needs plans. March 30.

Centers for Medicare & Medicaid Services, Department of Health and Human Services. 2009c. *Roadmap for quality measurement in the traditional Medicare fee-for-service program.* Baltimore, MD: CMS.

Centers for Medicare & Medicaid Services, Department of Health and Human Services. 2010. Medicare and Medicaid programs; electronic health record incentive program. Proposed rule. *Federal Register* 75, no. 8 (January 13): 1844–2011.

Congressional Budget Office. 2009. Letter to the Honorable Charles B. Rangel concerning the effect on federal direct spending and revenues of the Health Information Technology for Economic and Clinical Health (HITECH) Act. January 21.

Congressional Research Service. 2009. *The Health Information Technology for Economic and Clinical Health (HITECH) Act*. February 23.

Eicheldinger, C., and A. Bonito. 2008. More accurate racial and ethnic codes for Medicare administrative data. *Health Care Financing Review* 29, no. 3 (Spring): 27–42.

Federman, A., and A. Siu. 2004. The challenge of studying the effects of managed care as managed care evolves. *Health Services Research* 39, no. 1 (February): 7–12.

Fisher, E. S., D. C. Goodman, and A. Chandra. 2008. Disparities in health and health care among Medicare beneficiaries. A Dartmouth Atlas Project report commissioned for the Robert Wood Johnson Foundation Aligning Forces for Quality program. December.

Harvard Medical School. 2009. The role of private plans in Medicare. Department of Health Care Policy description of research agenda. http://www.hcp.med.harvard.edu/research/projects/newhouse_Medicare_PO1.

Hayward, R. A. 2008. Access to clinically-detailed patient information: A fundamental element for improving the efficiency and quality of healthcare. *Medical Care* 46, no. 3 (March): 229–231.

Health Information Technology Policy Committee. 2009. *Health IT Policy Committee recommendations to national coordinator for defining meaningful use: Final.* Washington, DC: Health Information Technology Policy Committee. August.

Heidenreich, P. A., M. McClellan, C. Frances, et al. 2002. The relation between managed care market share and the treatment of elderly fee-for-service patients with myocardial infarction. *American Journal of Medicine* 112, no. 3 (February):176–182.

Institute of Medicine. 2006. *Quality measurement: Accelerating improvement.* Washington, DC: IOM.

Keenan, P. S., M. N. Elliott, P. D. Cleary, et al. 2009. Quality assessments by sick and healthy beneficiaries in traditional Medicare and Medicare managed care. *Medical Care* 47, no. 8 (August): 882–888.

Kmetik, K. S., J. Chung, and S. Sims. 2007. The quality of quality measures. *American Journal of Managed Care* 13, no. 10 (October): 547–549.

Landon, B. E., A. M. Zaslavsky, S. L. Bernard, et al. 2004. Comparison of performance of traditional Medicare vs. Medicare managed care. *Journal of the American Medical Association* 291, no. 14 (April 14): 1744–1752.

McBean, A., Z. Huang, B. Virnig, et al. 2003. Racial variation in the control of diabetes among elderly Medicare managed care beneficiaries. *Diabetes Care* 26, no. 12: 3250–3256.

Medicare Payment Advisory Commission. 2004. *Report to the Congress: Medicare payment policy*. Washington, DC: MedPAC.

Medicare Payment Advisory Commission. 2005a. *Report to the Congress: Medicare payment policy*. Washington, DC: MedPAC.

Medicare Payment Advisory Commission. 2005b. *Report to the Congress: Issues in a modernized Medicare program*. Washington, DC: MedPAC.

Medicare Payment Advisory Commission. 2007. *Report to the Congress: Promoting greater efficiency in Medicare*. Washington, DC: MedPAC.

Medicare Payment Advisory Commission. 2008. *Report to the Congress: Medicare payment policy*. Washington, DC: MedPAC.

Medicare Payment Advisory Commission. 2009a. Comment letter on the draft recommendations presented by the Meaningful Use Workgroup of the HIT Policy Committee. Washington, DC: MedPAC. June 26.

Medicare Payment Advisory Commission. 2009b. *Report to the Congress: Improving incentives in the Medicare program.* Washington, DC: MedPAC.

National Committee for Quality Assurance. 2009a. E-mail message to Medicare Payment Advisory Commission. December 28.

National Committee for Quality Assurance. 2009b. Personal communication with Phyllis Torda, senior executive for strategic initiatives.

National Committee for Quality Assurance. 2009c. *Special needs plans structure and process measures (revised)*. CMS Contract No. HHSM–500–2006–00060C. March.

National Quality Forum. 2008. *Recommended common data types and prioritized quality measures for electronic healthcare information systems*. Washington, DC: NQF.

National Quality Forum. 2009. Memorandum to NQF members regarding pre-voting review for National Voluntary Consensus Standards for Hospital Care: Outcomes and Efficiency, Phase II. April 22.

New York University Center for Health and Public Service Research. 2009. *NYU ED algorithm background*. New York: New York University Center for Health and Public Service Research. http://wagner.nyu.edu/chpsr/.

Shahian, D. M., T. Silverstein, A. F. Lovett, et al. 2007. Comparison of clinical and administrative data sources for hospital coronary artery bypass graft surgery report cards. *Circulation* 115 (March 27): 1518–1527.

Trivedi, A. N, W. Rakowski, and J. Ayanian. 2008. Effect of cost sharing on screening mammography in Medicare health plans. *New England Journal of Medicine* 358 (January 24): 4.

Wenger, N. S., C. P. Roth, P. Shekelle, et al. 2007. Introduction to the assessing care of vulnerable elders–3 quality indicator measurement set. *Journal of the American Geriatrics Society* 55 (October): S247–S252.