Applying the Commission's principles for measuring quality: Population-based measures and hospital quality incentives

Applying the Commission's principles for measuring quality: Population-based measures and hospital quality incentives

Chapter summary

The Commission has recommended that Medicare link payment to quality of care to reward accountable entities and providers for offering high-quality care to beneficiaries. The Commission has recently formalized a set of principles for measuring quality in the Medicare program. Overall, quality measurement should be patient oriented, encourage coordination, and promote delivery system change. Medicare quality incentive programs should use a small set of outcomes, patient experience, and value measures to assess the quality of care across different populations, such as beneficiaries enrolled in Medicare Advantage (MA) plans, accountable care organizations (ACOs), and fee-for-service (FFS) in defined market areas, as well as those cared for by specified hospitals, groups of clinicians, and other providers. Applying the Commission's principles, Medicare quality incentive programs should score these risk-adjusted, population-based measure results against absolute performance thresholds and then use peer grouping to determine payment adjustments based on the provider's quality performance. In this chapter, we first apply the Commission's principles to two population-based outcome measures (potentially preventable admissions and home and community days (formerly known as "healthy days at home")) that may be used to evaluate quality of care for different populations. Next, we apply the principles to the design of a new hospital quality incentive program that combines measures of hospital outcomes, patient experience, and Medicare spending per beneficiary.

In this chapter

Applying the Commission's principles for measuring quality to population-based measures

We analyzed the utility of two population-based measure concepts to assess the quality of FFS care at market-area levels (e.g., geographic areas representing local health care market areas) and whether there is enough variation in performance to allow comparisons of FFS quality of care across market areas.

Potentially preventable admissions

Potentially preventable admissions (PPAs) constitute an important quality measure because hospitalizations for conditions such as diabetes and pneumonia can potentially be preventable if ambulatory care is provided in a timely and effective manner. To build on the Commission's work testing the measurement of PPAs in FFS Medicare and across Medicare payment models, we applied a quality measure developed for MA to FFS administrative claims data.

We calculated the observed rate of PPAs per 1,000 FFS beneficiaries for both chronic (e.g., diabetes) and acute (e.g., bacterial pneumonia) conditions. We found that observed (that is, not risk-adjusted) PPA rates varied across population groups (e.g., age, sex, Medicaid eligibility) and across two different definitions of market areas. This variation signals opportunities to improve the quality of care within areas and the potential to use the measure to compare quality across local health care markets. However, more development is needed to incorporate risk adjustment based on FFS data in the analysis.

Home and community days

The Commission tested a prototype home and community days (HCDs) measure to assess how well health care markets and organizations that take responsibility for a population keep people alive and out of health care institutions. The HCD measure is defined as 365 days minus the sum of (1) the number of days in the year that a beneficiary spends in certain institutional (e.g., hospital, skilled nursing facility) and ambulatory (e.g., emergency department) health care settings and (2) the number of mortality days (i.e., the number of days in the year that a beneficiary was not living, if any).

We calculated risk-adjusted HCDs from 2013 to 2015 for two populations of FFS beneficiaries (all beneficiaries 65 years and older and beneficiaries 65 years and older with two or more chronic conditions). In 2015, the adjusted HCD rate for beneficiaries 65 years and older was 351 days compared with 328 days for beneficiaries 65 years and older with 2 or more chronic conditions. We also compared the distribution of mean, risk-adjusted HCDs by MedPAC-defined market areas and hospital service areas. For the group of all beneficiaries 65

years and older, the difference in HCDs was only 3 days between high- and low-performing market areas; the difference was only 9 days for the group with 2 or more chronic conditions. However, because of the limited variation in HCDs over market areas and the challenges posed by the need to develop appropriate weights for constructing the composite measure, the Commission questions the immediate utility of the HCD measure in its current form to assess market-level FFS performance.

The Commission has continued interest in developing claims-calculated, population-based outcome measures. Ideas for population-based measures include "mean time between failure" (e.g., mean time between hospitalizations), successful community discharge, home-to-home transition time, end-of-life care and burdensome transitions, and low-value care.

Applying the Commission's principles for measuring quality to hospital quality incentives

We also examined the potential to create a single quality-based payment program for hospitals in light of Medicare's experience with four hospital payment incentive programs: the Hospital Inpatient Quality Reporting Program, Hospital Readmissions Reduction Program, Hospital-Acquired Condition Reduction Program (HACRP), and Hospital Value-based Purchasing (VBP) Program. The Commission's and others' main concerns about these programs are that (1) there are too many overlapping hospital quality payment and reporting programs, which creates unneeded complexity in the Medicare program; (2) all-condition measures are more appropriate to measure the performance of hospitals rather than the condition-specific readmissions and mortality measures currently used; (3) the existing programs include process measures and measures not consistently reported by providers; and (4) some of the programs score hospitals using "tournament models" (providers are scored relative to one another) rather than on clear, absolute, and prospectively set performance targets.

Ideally, the Congress could redesign the multiple hospital quality payment programs under a single hospital value incentive program (HVIP) that would be patient oriented, encourage coordination across providers and time, and promote change in the delivery system. Although CMS likely has the authority to make some of our suggested changes to hospital quality payment without congressional action (e.g., improving public reporting), other key reforms would require statutory changes.

The Commission asserts that the Medicare program should consider differences in providers' patient populations—which affect providers' performance on quality measures, including social risk factors—and that Medicare should account for

social risk factors in quality programs by adjusting payment through peer grouping. Applying these principles, we modeled an HVIP in which quality-based payments are distributed to hospitals organized under 10 peer groups based on the share of fully dual-eligible beneficiaries treated. (Fully dual-eligible beneficiaries are covered by both Medicare and Medicaid, and so we use this population category as a proxy for low income as a social risk factor.) In our model, the HVIP is budget neutral, with awards funded by a payment withhold from all hospitals. Our HVIP model uses a 2 percent withhold, which is the same as the existing VBP program uses, but policymakers could raise or lower the withhold amount.

Under our HVIP model, relative to the 2 percent withhold, about half of hospitals would receive a negative payment adjustment, and about half would receive a positive adjustment. Most hospitals rewarded under the existing programs would continue to receive rewards, and hospitals currently incurring penalties would continue to do so. Our peer grouping of hospitals allowed us to examine how hospitals serving large shares of low-income patients perform. We found that, compared with the existing quality payment programs, the HVIP approach makes payment adjustments among hospitals that serve different populations more equitable. Over the next year, the Commission plans to continue to refine a design for an HVIP that conforms with our principles for quality measurement. Some topics the Commission will further explore include weighting of measures, withhold values, patient experience measures, and patient safety measures. ∎

Introduction

The Commission contends that Medicare payments should not be made without consideration of the quality of care delivered to beneficiaries. The Congress has enacted quality reporting programs for almost all of the major fee-for-service (FFS) provider types and for Medicare Advantage (MA) and Part D plans, and it has mandated pay-for-performance (which Medicare refers to as value-based purchasing) for hospitals, dialysis facilities, physicians, accountable care organizations (ACOs), and skilled nursing facilities. Over the past several years, the Commission has expressed concern that Medicare's quality measurement programs are "overbuilt," relying on too many clinical process measures that are, at best, weakly correlated with health outcomes of importance to beneficiaries and the program. Relying on a large number of process measures can reinforce payment incentives in FFS to overprovide and overuse measured services. Process measures are also burdensome for providers to report, while yielding limited information to support clinical improvement. Although CMS has been shifting away from process to outcome measures in some of the Medicare quality programs, more work is needed to align the quality measurement systems with the Commission's principles for measuring quality (see text box, p. 180).

Applying quality measurement principles across populations

In the June 2014 and 2015 reports to the Congress, the Commission put forth a concept for an alternative to Medicare's current system for measuring the quality of care provided to beneficiaries (Medicare Payment Advisory Commission 2015a, Medicare Payment Advisory Commission 2014). This alternative led to the development of the Commission's principles on quality measurement—in particular, encouraging providers to work across the delivery system. Under this alternative policy, Medicare quality incentive programs would use a small set of outcomes, patient experience, and value measures to assess the quality of care across different populations, such as beneficiaries enrolled in Medicare Advantage (MA) plans, accountable care organizations (ACOs), and fee-for-service (FFS) in defined market areas, as well as those cared for by specified hospitals, groups of clinicians, and other providers. Medicare can link quality performance to payment using such measures to create incentives for MA plans, ACOs, and providers to offer high-quality care

to beneficiaries. Based on the Commission's principles, Medicare quality incentive programs for these accountable entities should score risk-adjusted measure results against absolute performance thresholds and then use peer grouping to adjust payment based on performance. Medicare's use of the same set of measures and scoring framework across populations could also promote other payers (e.g., Medicaid and commercial) using the same systems, which could reduce the burden providers face in tracking a diverse number of quality measures and methodologies across payers.

In this chapter, we first apply the Commission's alternative policy and principles to test the use of two population-based outcome measures (potentially preventable admissions (PPAs) and home and community days (HCDs) (formerly known as "healthy days at home")) to evaluate FFS quality of care and beneficiary access to health care in local health care market areas. We wanted to test the use of the measures for the FFS population in health care markets before applying the measures to other populations. Next, we apply the Commission's principles to the design of a new hospital quality payment program that uses current hospital outcome, patient experience, and Medicare spending per beneficiary measures.

Applying the Commission's principles for measuring quality to population-based measures

This chapter presents our analysis of two claims-based, population-based measures: PPAs and HCDs. Our analyses are meant to test whether the two measures can be used to evaluate quality of care for FFS beneficiaries and compare performance across local health care market areas, before applying the measures to other populations.

Potentially preventable admissions

Hospital stays can pose risks to patients, particularly the elderly. Adverse events represent a prominent risk, including iatrogenic infections, medication errors, device failures, and pressure injuries such as decubitus ulcers. According to researchers at the Centers for Disease Control and Prevention (CDC), on any given day, approximately 1 in 25 U.S. patients contracts at least one infection during the course of hospital care

The Commission has recently formalized a set of principles for measuring quality in the Medicare program, principles that we apply in developing quality measures, modeling the design or redesign of quality incentive (or value-based purchasing) programs, and commenting on CMS proposals for quality measurement. Over recent years, the Commission has articulated elements of these principles in its policy development process, but we now present them in a complete framework for evaluating Medicare's approaches to assessing quality of care. The Commission's principles are as follows:

- Quality measurement should be patient oriented, encourage coordination across providers and time, and promote relevant change in the nature of the delivery system.

- Quality measurement should not be unduly burdensome for providers.

- Medicare quality programs should include population-based measures such as outcomes, patient experience, and value (e.g., Medicare spending per beneficiary, measures of services that have little or no clinical benefit). Providers may choose to use more granular measures to manage their own quality improvement.

- Medicare quality programs should give rewards based on clear, absolute, and prospectively set performance targets (as opposed to "tournament models," under which providers are scored relative to one another).

- The Medicare program should take into account, as necessary, differences in a provider's patient population, including social risk factors. Because adjusting measure results for social risk factors can mask disparities in clinical performance, Medicare should account for social risk factors by directly adjusting payment through peer grouping.

- Medicare should target technical assistance resources to low-performing providers.

- Medicare should support research and data collection to reduce measurement bias, including, for example, the effects of social risk factors.

The Commission also maintains that the goal of improved care should extend to all patients, regardless of health status, income, and race. Recognizing that those expectations are more likely to be met if they are combined with additional resources to accelerate a provider's ability to address particularly challenging care delivery environments, the Commission recommended in June 2011 that the Quality Improvement Organization Program be fundamentally restructured and that funding be reprogrammed to give providers and communities more choices in who assists them in quality improvement activities and flexibility in how resources can be used. (Medicare Payment Advisory Commission 2011). The Commission also recommended that Medicare make technical assistance to low-performing providers and community initiatives a high priority as a strategy to complement payment policy and address persistent health care disparities. ∎

(Centers for Disease Control and Prevention 2016b). In addition, the inpatient environment itself can lead to a reduction in elderly patients' independence as they cope with functional loss that can stem from extended bed rest (Covinsky et al. 2011). Furthermore, the hospital environment often hinders discussion about treatment options.

Hospitalizations due to conditions such as diabetes and pneumonia are potentially preventable if ambulatory care is provided in a timely and effective manner.[1] PPAs can fall into five categories: system related (e.g., unavailability of services), physician related (e.g., suboptimal monitoring), medical (e.g., medication side effects), patient related (e.g., delay in seeking help), and social (e.g., lack of social support) (Freund et al. 2013). Evidence also suggests that effective primary care is associated with lower PPAs (Gao et al. 2014). The patient may have required acute-level services at the time he or

she sought care, but the need for the admission might have been avoided with appropriate ambulatory care and coordination activities.

Rates of PPAs calculated through administrative claims data can reflect the quality of the care provided under payment models and by providers in a local market area (that is, a defined population). High-quality MA plans in a local market area should be able to manage beneficiary, hospital, and physician relations to coordinate care and provide appropriate access (Wholey et al. 2003). High-quality ACOs should also be able to manage relationships to improve care. For example, ACOs can provide tools and data to clinicians about patients with chronic ambulatory care–sensitive conditions (such as diabetes and asthma) so they can appropriately monitor, coordinate, and follow up with patients and reduce avoidable hospitalizations. FFS clinicians can also play a role in affecting admissions in the ambulatory care area they serve by effectively coordinating with other providers and offering adequate access to beneficiaries. For example, a clinician's availability for appointments can affect how well a patient's chronic conditions are managed and whether a patient's acute conditions (such as pneumonia) can be identified and treated outside of the hospital in a timely manner (Davies et al. 2009).

The Commission's prior work on measuring PPAs

In our June 2014 and 2015 reports to the Congress, the Commission included PPAs for inpatient hospital care as a population-based measure concept for evaluating quality in a market area (Medicare Payment Advisory Commission 2015a, Medicare Payment Advisory Commission 2014). Over the years, we have used two existing measure specifications to define the concept of "potentially preventable" and measure FFS rates accordingly. In the March 2017 report to the Congress, we presented national rates and variation by market areas using a definition of PPAs developed by 3M (Medicare Payment Advisory Commission 2017). During the Commission's November 2016 meeting, the Commission expressed concern that the 3M[TM] measure was not available in the public domain and that providers could find the measure definitions overly complicated. In recent March reports to the Congress, we also published CMS-reported rates of hospitalizations based on the Agency for Healthcare Research and Quality (AHRQ) Prevention Quality Indicator (PQI) specifications for three individual ambulatory care–sensitive conditions (diabetes, congestive heart failure (CHF), and bacterial pneumonia).

Calculating potentially preventable admissions in the FFS population

To further test the concept of measuring PPAs for FFS beneficiaries and to compare performance across market areas, we used a 2018 measure specification developed by AHRQ and adopted with permission by the National Committee for Quality Assurance (NCQA). The measure specification is publicly available as part of NCQA's Healthcare Effectiveness Data and Information Set[®] (HEDIS[®]), and the measure is written for the Medicare population, specifically for MA plans to report.[2] In the summer of 2018, MA plans will report measure results to CMS using the 2018 measure specification (collected by NCQA), along with other quality measures that are used to calculate star ratings.[3] Thus, in the future, we may have the ability to use one PPA measure specification to compare performance across MA, FFS, and ACOs nationally and within markets.

The HEDIS (MA) PPA measure represents the observed rate of PPAs and the risk-adjusted ratio of observed-to-expected potentially preventable admissions. PPAs are calculated for chronic conditions (e.g., diabetes) and acute conditions (e.g., pneumonia).[4,5] (Although we chose to analyze this measure specification for FFS, we are not endorsing any approach to measuring PPAs. We are simply exploring the use of PPAs as a population-based measure of ambulatory care.) Comparing FFS and MA plan quality performance in a local area is a future goal of this work, so we did not make changes to the HEDIS specification in order to permit "apples-to-apples" comparisons among Medicare payment models. Our analysis examines PPAs for Medicare FFS beneficiaries ages 67 and older because the HEDIS specification requires two years of beneficiary enrollment in the MA plan. In future analyses, we could apply the measure to different populations, including the under–age 67 population. We did not calculate risk-adjusted numbers of expected discharges because the regression model NCQA uses to calculate the expected results is based on the risk profiles of a sample of MA beneficiaries. Since MA plan populations and the coding intensity of diagnoses differ from FFS, we would need to develop FFS-based risk weights to calculate expected results. We therefore focused our analysis on the observed rate of unadjusted PPAs per 1,000 beneficiaries ages 67 and older. We also focused on national results and not results at the market level because unadjusted results would not capture any underlying differences in market-area population characteristics.

**TABLE
7–1**

PPAs per 1,000 FFS beneficiaries ages 67 and older vary by population group, 2016

	PPA rate per 1,000 FFS beneficiaries		
	Acute conditions	**Chronic conditions**	**Total**
Age groups			
67–74	8.4	12.0	20.3
75–84	17.0	20.0	37.0
85+	34.3	31.3	65.6
Sex			
Male	13.2	17.6	30.9
Female	17.0	17.8	34.8
Medicaid eligibility			
Fully dual eligible	32.2	34.0	66.1
Partially dual eligible	22.4	33.3	55.7
Non–dual eligible (Medicare only)	13.3	15.5	28.7
Total	15.3	17.7	33.1

Note: PPA (potentially preventable admission), FFS (fee-for-service). To evaluate the utility of measuring PPAs for FFS beneficiaries, we calculated the observed (not risk-adjusted) rates of admissions tied to acute (e.g., pneumonia), chronic (e.g., diabetes) and total (acute plus chronic) conditions. Rates presented are the number of PPAs divided by the number of beneficiaries in the qualifying population, multiplied by 1,000. The qualifying population is the same across the acute and chronic categories. Beneficiaries who died in the measurement year are excluded. Fully dual-eligible beneficiaries qualify for a full range of Medicaid benefits, and partially dual-eligible beneficiaries qualify for Medicaid payment of the Medicare premium and perhaps the cost sharing for Medicare services.

Source: MedPAC analysis of 2016 Medicare claims data.

We found that it is feasible to calculate unadjusted, observed PPAs for FFS beneficiaries nationally and for two different geographic area levels representing local health care markets (MedPAC-defined market areas designed to match insurance markets served by private plans and Dartmouth-defined hospital service areas (HSAs), which are collections of zip codes that represent a local market area whose residents receive most of their inpatient care from the hospitals in that area). We also found variation by population groups (e.g., age, sex, Medicaid eligibility) and by market area, which signals both opportunities to improve quality performance within areas and the measure's potential for comparing quality across local health care markets.

In the future, the Commission could develop a risk adjustment model to calculate FFS and ACO expected PPA rates and compare market-area risk-adjusted PPAs. The risk adjustment model would need to ensure that the PPA measure primarily reflects an organization's or area's quality of care rather than underlying differences in patient severity. Using the MA PPA measure as an example, we can test risk adjustment using age, sex, and disease

severity based on CMS's hierarchical condition categories (CMS–HCCs) because we have access to these FFS data. In the future, if the PPA measure is considered for a Medicare quality payment program, we can test the use of peer grouping to account for differences in the social risk factors of populations. The Commission continues to encourage CMS to support research and data collection to improve our ability to take into account social risk factors.

Qualifying population   The qualifying population for the PPA measure is all FFS beneficiaries who meet the following criteria: are ages 67 years and older at the end of the measurement year, are alive at the end of the measurement year, are continuously enrolled in Part A and Part B for the measurement year and the previous year with no months of MA enrollment, and have used no hospice services in the measurement year. For the 2016 measurement year, the population of FFS beneficiaries who met those criteria was about 22.5 million nationwide.

Beneficiaries with three or more discharges in the measurement year were considered outliers and removed from the qualifying population and observed event

| | PPA rate per 1,000 FFS beneficiaries | | |
	Acute conditions	Chronic conditions	Total
National mean	15.3	17.7	33.1
10th percentile (highest performing)	10.4	11.1	22.4
50th (median)	16.2	17.8	34.9
90th (lowest performing)	24.3	24.9	48.7
Ratio of 90th to 10th percentile	2.3	2.2	2.2

Note: PPA (potentially preventable admission), FFS (fee-for-service). To evaluate the utility of measuring PPAs for FFS beneficiaries, we calculated the observed (not risk-adjusted) rates of admissions tied to acute (e.g., pneumonia), chronic (e.g., diabetes) and total (acute plus chronic) conditions. Rates presented are the number of PPAs divided by the number of beneficiaries in the qualifying population, multiplied by 1,000. The qualifying population is the same across the acute and chronic categories. Beneficiaries who died in the measurement year are excluded. There are over 1,200 MedPAC-defined market areas designed to match insurance markets served by private plans. The average qualifying population in each market area is about 19,000 beneficiaries.

Source: MedPAC analysis of 2016 fee-for-service Medicare claims data.

counts. Almost 57,000 outlier beneficiaries were removed from our measure calculation (about 75 percent had hospitalizations tied to chronic conditions vs. acute primary diagnoses).

PPAs by chronic and acute conditions As with the MA plan PPA measure, we calculated the number of inpatient admissions and observation stays tied to the beneficiaries in the qualifying population (both are observed events). The observed events include admissions with the primary diagnosis of the following *chronic* conditions: diabetes (short-term and long-term complications, uncontrolled diabetes, lower extremity amputation among patients with diabetes); chronic obstructive pulmonary disease (COPD); asthma; hypertension; and heart failure. Observed events also include admissions tied to beneficiaries with the primary diagnosis of the following *acute* conditions: bacterial pneumonia, urinary tract infection, cellulitis, and pressure ulcers. We calculated a total number of PPAs (chronic plus acute).

National PPAs results In 2016, PPAs accounted for about 8 percent of FFS Medicare hospital admissions.[6] Nationally, there were 15.3 acute-condition-related PPAs per 1,000 FFS beneficiaries and 17.7 chronic-condition-related PPAs per 1,000 FFS beneficiaries, for a total of 33.1 PPAs per 1,000 FFS beneficiaries (Table 7-1).

In 2016, older Medicare beneficiaries had higher PPA rates for both acute and chronic conditions. Female beneficiaries had higher PPA rates than men for acute conditions and about the same PPA rate as men for chronic conditions. Both fully (i.e., receive full range of Medicaid benefits) and partially (i.e., Medicaid pays Medicare premium and may also pay the cost sharing for Medicare services) dual-eligible beneficiaries had higher PPA rates for both acute and chronic conditions compared with non-dual-eligible (Medicare-only) beneficiaries. These patterns are consistent with CMS-produced results using selected AHRQ PQIs and with our prior work using the 3M PPA measure.

The pattern of higher PPA rates for the dual-eligible population is also expected when comparing admission rates that are not risk adjusted for population characteristics. For example, the fully dual-eligible population is older than the partially dual-eligible population, which may explain the fully dual-eligible population's higher rate of PPAs. In future analyses of PPA rates, we will consider the effect of dual eligibility on the PPA results.

Distribution of PPAs in local health care market areas Differences in PPA results across local health care markets can help distinguish differences in quality compared with a national mean and across market areas. We calculated PPA rates for acute and chronic conditions and total PPAs per 1,000 FFS beneficiaries in the 1,200 MedPAC market areas that the Commission recommends for MA payment and quality reporting (Table 7-2). We found that

total observed (not risk-adjusted) PPA rates varied across market areas, with the market area in the 90th percentile of PPA rates having a rate that was 2.2 times the market area in the10th percentile. The magnitude of difference between the market areas in the 90th and 10th percentiles of observed PPA rates for acute conditions and chronic conditions individually was similar to that for the total PPA rate.

To model rates at a more narrowly defined health care market level (that is, the Dartmouth-defined hospital service areas (HSAs)), we calculated PPA rates for acute and chronic conditions and total PPA rates per 1,000 FFS beneficiaries in the roughly 3,400 HSAs. An HSA is a collection of zip codes that represents a local market area whose residents receive most of their inpatient care from the hospitals in that area. As with the larger MedPAC market areas presented in Table 7-2, PPA rates varied across HSAs, with HSAs in the 90th percentile of PPA rates exceeding HSAs in the 10th percentile of PPA rates by 2.1 times (data not shown). PPA rates for acute conditions had slightly more variation compared with PPA rates for chronic conditions.

Home and community days measure

The Commission tested a "home and community days" (HCDs) quality measure to assess how well health care organizations keep people healthy and out of health care institutions.[7] We chose to focus on the number of days per year that beneficiaries did not receive institutionalized medical care (such as days during which a beneficiary did not have an inpatient stay) and mortality days. An alternative to the measure could include days in which beneficiaries had any interaction with the health system (i.e., days in which Medicare covered any medically necessary service such as a physician office visit or an inpatient stay (Medicare Part A and Part B)).

High-quality MA plans and ACOs are designed to manage beneficiary, hospital, and physician relations to coordinate care and provide appropriate access to keep people out of health care institutions. For example, ACOs can provide tools and data to physicians about patients with ambulatory care–sensitive conditions (such as diabetes and asthma) so that they can appropriately monitor, coordinate, and follow up with patients and reduce inpatient stays. FFS clinicians can also play a role in affecting HCDs in their ambulatory care area by effectively coordinating with other providers and offering adequate access to beneficiaries.

Commission staff worked with a team from the Harvard School of Public Health to develop a prototype HCD measure. As described in the June 2015 report to the Congress, an HCD measure using Medicare claims data may be a meaningful gauge for comparing differences in health status across populations and be less complicated than other measures for beneficiaries, policymakers, and other stakeholders to understand (Medicare Payment Advisory Commission 2015a).

CMS is actively developing a quality measure for Medicare and Medicaid health plans and long-term services and support populations based on the Commission's HCD measure. CMS may submit the measure, currently named "days in the community," for endorsement by the National Quality Forum. Also, in 2016, the National Bureau of Economic Research released a working paper, *Healthy-time Measures of Health Outcomes and Healthcare Quality*, that describes some conceptual and empirical foundations of "healthy-time" measures of health care quality. Their analysis features the Commission's developing HCD measure and similar measures from other organizations.[8] The authors concluded that "the basic premises underlying this [the Commission's] measure's definition are conceptually sound and intuitively appealing; its use as a patient-centered outcome or care-quality indicator holds promise" (Burns and Mullahy 2016).

Calculating home and community days

The Commission's HCD measure, for the purposes of this chapter, pertains to FFS Medicare beneficiaries 65 years and older, excluding those enrolled in MA for any part of the year and those not enrolled in Medicare FFS continuously throughout the year. For the HCD measure we modeled, we focused on beneficiary interactions with more serious health care that is covered by Medicare and on mortality. We defined this measure algorithmically as follows:

> *HCDs = 365 days – (days in short-term acute care hospital + days in inpatient rehabilitation facility + days in long-term care hospital + days in inpatient psychiatric facility + days in skilled nursing facility + days in observation status + days of emergency department use + mortality days)*

For each FFS beneficiary, we calculated his or her total number of mortality days, which is defined by the number of days remaining in the calendar year after the date of death. For example, a beneficiary who did not die during

the year would have no mortality days. A beneficiary who died on December 28 would have three mortality days for the year. Inpatient, observation, skilled nursing facility (SNF), inpatient psychiatry, inpatient rehabilitation, and long-term care hospital days were defined as the total number of days per year the beneficiary spent in each of these respective settings. For the purposes of this analysis, we weighted HCD components equally, but policymakers interested in developing this measure further could give the components different weights based on some prioritization that takes into account interests shared by the Medicare program and its beneficiaries.

We did not subtract home health visit days in calculating a beneficiary's HCDs. Home health represents a midpoint at which the patient is at home but is still in need of health care services. In some health care markets, home health visits are used to prevent or limit use of other, more expensive services—in particular, inpatient and SNF care. Subtracting home health visit days from the HCD measure could therefore penalize these markets and providers unfairly. Documented overuse of home health care could make a case for subtracting home health visits from the HCD measure. For instance, the Office of Inspector General (OIG) has recently identified 27 geographic areas as "hotspots" for characteristics commonly found in OIG-investigated cases of home health fraud, so, in some markets, penalizing home health use could be an appropriate approach (Office of Inspector General 2016). Yet even with these potential differences in home health use by market area, from the beneficiary's perspective, home health visits are likely more desirable than the use of other health care services that would lower HCDs, a circumstance that argues for not subtracting home health visit days from the HCD measure.

Risk adjustment modeling

A critical step in the development of the HCD measure is to test appropriate risk adjustment models. Such models should ensure that the HCD measure primarily reflects an organization's or area's quality of care rather than underlying differences in patient severity. Using linear regression, we developed a model that included variables readily available in FFS claims data and used in other quality measures: age, sex, disease burden determined from HCCs, and market-fixed effects (e.g., local characteristics).[9]

We found that disease burden had the greatest impact on HCDs. The diseases or conditions that had the most effect on HCDs were respiratory arrest, nephritis, extensive third-degree burns, seizure disorders and convulsions, and coma/brain compression/anoxic damage (all statistically significant). Our analysis found that HCDs decrease with age. Men had slightly more HCDs than women.

Effect of dual-eligibility status We also tested the effects that social risk factors could have on the risk adjustment model. In a separate regression model, we included race and dual eligibility (defined by a beneficiary having both Medicare and Medicaid coverage for at least one month of a year). (Dual eligibility may be a proxy for low income.) When included as a variable, dual-eligibility status had some impact on HCDs (regression coefficient = –7.76) (i.e., dual-eligible status corresponds with fewer HCDs). Coefficients for race were not significant. When we compared the explanatory power (R^2) of a risk adjustment model with age, sex, disease burden, and market-fixed effects with a model that included those variables plus dual-eligibility status, there was no difference in the explanatory power of the models (both $R^2 = 0.32$).

Since dual eligibility seemed to have some impact for individual beneficiaries but not on the overall model's explanatory power, we investigated the impact of dual-eligibility on market-area performance.[10] We examined how market performance varies among high-share versus low-share dual-eligible markets and found that mean HCDs decline with increasing deciles of dual-eligible share, although the relationship is not constant (Figure 7-1, p. 186). Among all beneficiaries ages 65 and older, markets in the top decile of dual-eligible share—in which more than 37 percent of beneficiaries were Medicaid eligible—had, on average, about 4 fewer HCDs compared with markets in the bottom decile of dual-eligible share (in which over 9 percent of beneficiaries were dual eligible). Among beneficiaries ages 65 and older with 2 or more chronic conditions, markets in the top decile of dual-eligible share had, on average, about 6 fewer HCDs compared with markets in the bottom decile of dual-eligible shares.[11]

We found that mortality days tended to be somewhat higher in markets with high dual-eligibility shares, resulting in a lower average number of HCDs, although the differences were small. Inpatient and SNF days were stable across the deciles of markets.

If CMS opted to use HCDs to compare quality across market areas or providers, the Secretary should be

FIGURE
7–1

Mean adjusted home and community days are slightly lower in local health care market areas with higher shares of dual-eligible beneficiaries, 2015
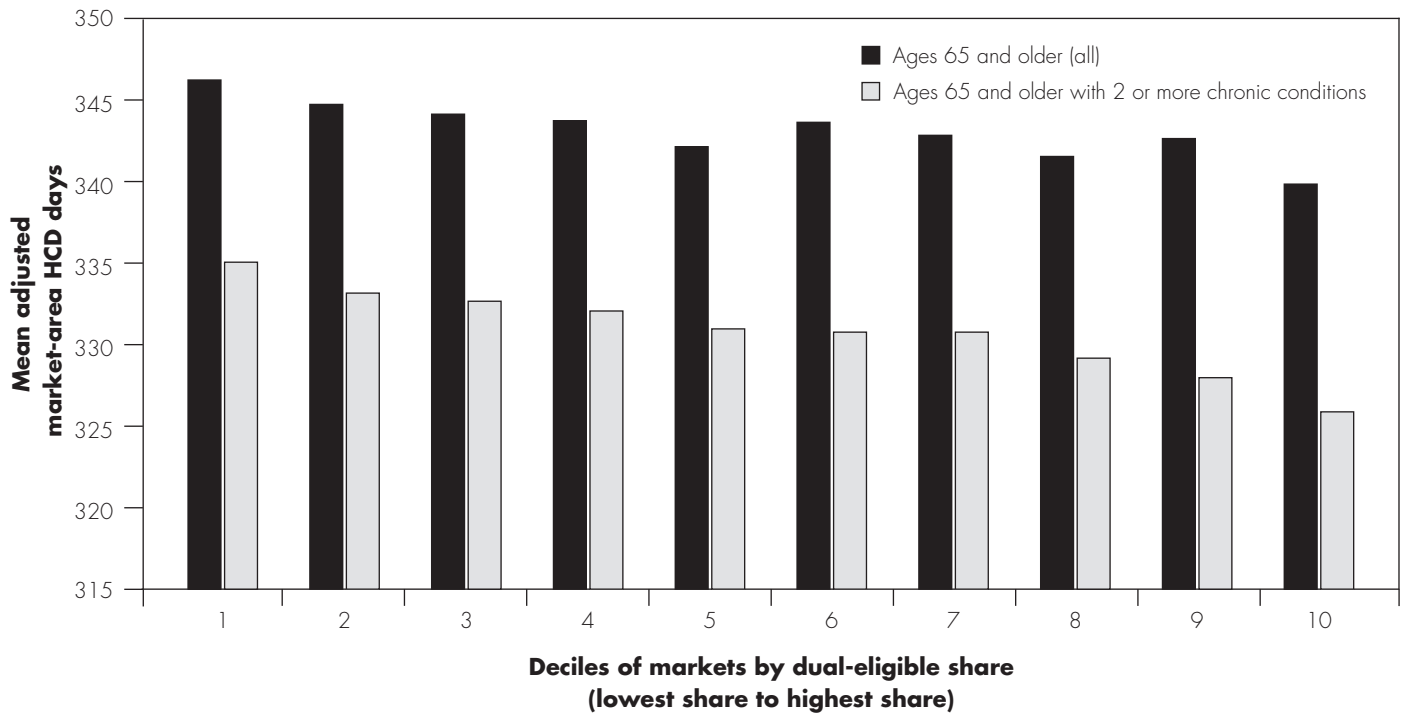
Note: Home and community days (HCDs) are adjusted for age, sex, and disease burden. There are over 1,200 MedPAC-defined market areas, designed to match insurance markets served by private plans. Deciles of markets were created based on the share of Medicare beneficiaries who were partially dual eligible for Medicare and Medicaid at any point in the year. Partially dual-eligible beneficiaries qualify for Medicaid payment of the Medicare premium and perhaps the cost sharing for Medicare services. Markets in Decile 1 have the lowest share of partially dual-eligible beneficiaries (9.4 percent), while markets in Decile 10 have the highest share of Medicaid-eligible beneficiaries (37.3 percent). The scale of the y-axis was chosen to highlight the differences in HCDs across market areas.

Source: Analysis of fee-for-service Medicare claims data, 2015.

cognizant of differences that correlate with dual eligibility. However, the Commission does not support the inclusion of dual-eligibility status in a risk adjustment model because doing so would mask disparities in clinical performance. Rather, Medicare should account for social risk factors by directly adjusting payment through peer grouping. The Commission continues to encourage the Secretary to support research and data collection to improve Medicare's ability to account for the effect of social risk factors on health outcomes.

Adjusted HCDs in local health care market areas

To understand HCDs for Medicare beneficiaries of different health status in different market areas over time, we calculated mean, risk-adjusted HCDs for the two different geographic area levels representing local health care markets that we used in the PPA analysis (MedPAC-

defined market areas designed to match insurance markets served by private plans and Dartmouth-defined HSAs, which are collections of zip codes that represent a local market area whose residents receive most of their inpatient care from the hospitals in that area). We calculated HCDs in each MedPAC-defined market area and HSA using 3 years of FFS Medicare data (2013 to 2015) for 2 populations: (1) all beneficiaries 65 years and older and (2) beneficiaries 65 years and older with at least 2 chronic conditions. There were at that time about 27.3 million beneficiaries 65 years and older, and about 7.7 million of those had at least two chronic conditions.

As expected, we found that Medicare beneficiaries with greater chronic-condition burden had fewer HCDs (Table 7-3). In 2015, the adjusted HCD rate for beneficiaries 65 years and older was 351 days compared with 328 days for beneficiaries 65 years and older with 2 or more chronic

TABLE
7–3

Mean adjusted home and community days for FFS beneficiaries were stable from 2013 to 2015

	Home and community days		
	2013	**2014**	**2015**
All beneficiaries 65 years and older	351	351	351
Beneficiaries 65 years and older with 2 or more chronic conditions	331	332	328

Note: FFS (fee-for-service). Home and community days are adjusted for age, sex, disease burden, and market-fixed effects. There are over 1,200 MedPAC-defined market areas, which are designed to match insurance markets served by private plans.

Source: Analysis of FFS Medicare claims data, 2013–2015.

conditions (a difference of 23 days). From 2013 to 2015, the results for beneficiaries 65 years and older were stable (351 days in each year), but the average HCDs declined slightly over the three years for beneficiaries with 2 or more chronic conditions (from 331 to 328 days).

For both population groups, the components of the HCD algorithm with the biggest impact on a market area's HCDs were mortality days, SNF days, and inpatient days (Table 7-4). For beneficiaries 65 years and older, the components were stable over time. There was somewhat more change from 2013 to 2015 in the HCD components for the beneficiaries 65 years and older with 2 or more chronic conditions. In the 2013 to 2015 period, the mortality days for that population increased by about 2.3

days; SNF days slightly increased over the three years (from 6.2 days to 6.6 days).[12]

Distribution of adjusted HCDs in local health care market areas

Because our goal was to compare FFS quality across health care markets and across different populations, we looked for variation in HCD measure results across both MedPAC-defined market areas and HSAs. We calculated the distribution of HCDs for all beneficiaries 65 years and older and for beneficiaries 65 years and older with 2 or more chronic conditions across MedPAC-defined market areas (Table 7-5, p. 188). The distribution among MedPAC-defined market-area HCDs for both populations was very small. The difference between the 90th and

TABLE
7–4

Home and community day components were stable from 2013 to 2015

	All beneficiaries 65 years and older			Beneficiaries 65 years and older with 2 or more chronic conditions		
Component (days)	**2013**	**2014**	**2015**	**2013**	**2014**	**2015**
Mortality	9.8	9.4	9.7	21.4	20.5	23.7
Skilled nursing facility	2.1	2.1	2.0	6.2	6.0	6.6
Inpatient	1.6	1.5	1.5	4.7	4.5	4.8

Note: Components are part of the home and community days (HCDs) calculation and represent mortality days and/or days in which beneficiaries have interactions with more serious health care. Mortality, skilled nursing facility, and inpatient days have the biggest impact on a market area's HCDs. HCDs are adjusted for age, sex, disease burden, and market-fixed effects. There are 1,200 MedPAC-defined market areas designed to match insurance markets served by private plans.

Source: Analysis of fee-for-service Medicare claims data, 2013–2015.

TABLE
7–5

Distribution of home and community days did not vary across local health care market areas, 2015

	Home and community days	
	All beneficiaries 65 years and older	**Beneficiaries 65 years and older with 2 or more chronic conditions**
National mean	351	328
10th percentile (lowest performing)	349	323
50th (median)	351	328
90th (highest performing)	352	332
Ratio of 90th to 10th percentile	1.01	1.03

Note: Home and community days are adjusted for age, sex, disease burden, and market-fixed effects. There are over 1,200 MedPAC-defined market areas designed to match insurance markets served by private plans.

Source: Analysis of fee-for-service Medicare claims data, 2015.

10th percentile MedPAC-defined market areas (of 1,200 MedPAC-defined market areas) for beneficiaries 65 years and older was only 3 days. The difference between the 90th and 10th percentile MedPAC-defined market areas for beneficiaries 65 years and older with 2 or more chronic conditions was only 9 days. For both populations, the highest performing MedPAC-defined market area's HCDs were always almost equal that of the lowest performing MedPAC-defined market area's HCDs. We found similar results when calculating HCDs for the more narrowly defined HSAs (data not shown).

With so little variation across local market areas and the challenges posed by the need to develop appropriate weights for constructing the composite measure, the Commission questions the immediate utility of the HCD measure in its current form to assess market-level FFS performance.

Future work on population-based quality measures

The Commission and policymakers may explore the following claims-calculated, population-based measures to assess Medicare quality for different defined populations (e.g., FFS populations associated with local market areas and beneficiaries served by MA plans, ACOs, hospitals, post-acute care (PAC) providers, or groups of clinicians). These measures are in line with the Commission's quality measurement principles to use population-based outcome measures that are patient-oriented, encourage coordination, and promote delivery system change.

- *Mean time between failure*—"Mean time between failure" is a commonly used engineering measure of predicted elapsed time between inherent failures of a mechanical or electronic system during normal system operation. Policymakers could consider how to apply this concept to measure quality of care for Medicare beneficiaries (for example, how many days between serious health care interactions (e.g., mean time between hospitalizations) for Medicare beneficiaries).

- *Successful community discharge*—The Improving Medicare Post-Acute Care Transformation Act of 2014 mandated that CMS develop quality measures for PAC providers. Responding to this mandate, CMS has developed measures for each PAC setting that assess whether PAC providers successfully discharge beneficiaries to the community (e.g., rate of beneficiaries discharged to the community who do not have an unplanned admission to a hospital within a set period of time). The Commission also currently calculates rates of discharge to the community for some individual PAC settings. Policymakers could consider measuring successful community discharge across all PAC providers for different populations.

- *Home-to-home transition time*—Home-to-home transition time is a measure that adds time spent in a PAC facility to time spent in the hospital to capture the full span of a hospitalization episode (Barnett et al.

2017). This measure is patient centered since patients are interested in when they can return home from all institutional care. Policymakers could explore the use of this measure to assess the home-to-home transition times for different populations.

- *End-of-life care and burdensome transitions—* Research has shown that a growing number of older adults in the United States are dying at home, but many continue to face multiple health care transitions to different care sites and receive aggressive inpatient care in their final days (Teno et al. 2013). Policymakers can consider developing a quality measure that assesses potentially burdensome transitions in the last days or weeks of life.

- *Low-value care—*For several years, the Commission has expressed concern that beneficiaries are receiving low-value care, or care that has little or no clinical benefit and that can potentially harm them (Medicare Payment Advisory Commission 2017, Medicare Payment Advisory Commission 2015b) (see also Chapter 10 in this report). The Commission has examined national FFS population rates for certain services and procedures that are considered low value. Policymakers should continue to explore measures of low-value care for different populations.

Applying the Commission's principles for measuring quality to hospital quality incentives

The Commission contends that Medicare payments should not be made without consideration of the quality of care delivered to beneficiaries and has recently formalized a set of principles for quality measurement in the Medicare program. For several years, the Medicare program has provided hospitals with incentive payments based on the quality of care they give to FFS beneficiaries (see text box on current hospital quality and value payment programs, pp. 190–191). The quality of hospital care has been improving over the years, which is partly due to these programs. However, the hospital industry has raised concerns that the designs of these programs are complex, overlap, and send different performance signals to hospitals. In addition, aspects of the programs do not align with the Commission's principles for measuring quality in the Medicare program.

Issues with current hospital quality and value programs

The Commission has four main concerns about the design of the current hospital quality programs. The first is that too many overlapping hospital quality payment and reporting programs create unneeded complexity for hospitals and the Medicare program itself (Medicare Payment Advisory Commission 2016a, Medicare Payment Advisory Commission 2016b). Some of the quality measures are scored in multiple programs. For fiscal years (FYs) 2020 and 2021, CMS has proposed to remove much of the duplication in quality measures across programs. For example, CMS would continue to use the hospital-acquired infection measures to assess performance in the Hospital-Acquired Conditions Reduction Program (HACRP) but would remove these measures from the Inpatient Quality Reporting Program (IQRP) and Hospital Value-based Purchasing (VBP) Program (Centers for Medicare & Medicaid Services 2018b).

Second, the Commission believes that all-condition mortality and readmissions measures are more appropriate to measure hospitals' performance, rather than the condition-specific (e.g., acute myocardial infarction) measures that are scored in the IQRP, VBP Program, and Hospital Readmissions Reduction Program (HRRP). Using all-condition measures would increase the number of observations and reduce the random variation that single-condition readmission rates face under current policy (Medicare Payment Advisory Commission 2013).

Third, the IQRP includes process measures that are not tied to outcomes and are burdensome to report (e.g., fibrinolytic therapy received within 30 minutes of hospital arrival). Also, providers may not be consistent in how they report some of the measures included in the IQRP, VBP Program, and HACRP (e.g., chart-abstracted measures and hospital-acquired infections). For FYs 2020 and 2021, CMS has proposed removing some chart-abstracted process measures, such as median time from emergency department (ED) arrival to ED departure for admitted ED patients, from the IQRP because the data collection and reporting costs outweigh the benefit of their continued use (Centers for Medicare & Medicaid Services 2018b).

Fourth, the VBP Program, HRRP, and HACRP score hospitals using "tournament models" (i.e., providers are scored relative to one another), not on clear, absolute, and prospectively set performance targets. For example, the HACRP's statutory design penalizes 25 percent of hospitals every year, even if all hospitals significantly reduce their HAC rates. The Commission's principles for quality measurement encourage Medicare quality programs to use fixed targets.

Current hospital quality and value payment programs

The Medicare program adjusts hospital payment based on four quality payment programs. One program adjusts payment based on whether a hospital reports quality measure results, and three programs adjust payment based on quality performance. Although not tied to payment, CMS's public reporting of hospital quality performance on the Hospital Compare website, including their star ratings, is another avenue for comparing acute care hospitals.

Reporting

The Hospital Inpatient Quality Reporting Program (IQRP) has been in place since fiscal year 2009 and is built on the preceding voluntary Hospital Quality Initiative. The IQRP reduces a hospital's annual market basket (the measure of inflation in costs of goods and services used by hospitals in treating Medicare patients) update by 2.0 percent if it does not successfully report quality measure data. In fiscal year 2018, nearly all inpatient hospitals met the IQRP requirement and will receive the full annual market basket update. There are 61 quality measures in the fiscal year 2020 program (based on coverage year 2018 performance). Hospitals report about half of those measures to CMS (e.g., patient experience surveys, health care–associated infections, medical record–abstracted measures such as average number of minutes before outpatients with chest pain or possible heart attack got an electrocardiogram, use of safe surgery checklist), while the other half are claims-based outcome (e.g., readmissions) or cost measures that CMS calculates. (CMS has proposed to remove 39 measures from the IQRP for fiscal years 2020 and 2021 (Centers for Medicare & Medicaid Services 2018b).)

Incentives for higher quality

Three programs adjust hospital payment based on how the hospital performs on quality results: the Hospital Readmissions Reduction Program (HRRP), the Hospital-Acquired Conditions Reduction Program (HACRP), and the Hospital Value-based Purchasing (VBP) Program.

Hospital Readmissions Reduction Program The HRRP was implemented in fiscal year 2013.[13] As a part of this program, hospitals that have excess Medicare readmissions over a three-year period for selected conditions have their inpatient prospective payment system (IPPS) payments reduced. In fiscal year 2018, the readmissions policy applies to six conditions (acute myocardial infarction (AMI), heart failure, pneumonia, chronic obstructive pulmonary disease, total hip and knee arthroplasty, and coronary artery bypass graft surgery). In 2018, the payment penalty is capped at 3 percent of a hospital's base diagnosis related group (DRG) payments per year. In 2018, about 80 percent of hospitals will have payments reduced because of higher than average readmissions for at least one condition. Total penalties will be about $556 million in 2018, or 0.5 percent of Medicare's total IPPS payments. Research has shown that readmission rates for AMI, heart failure, and pneumonia decreased more rapidly after the HRRP began and that improvement was most marked for hospitals with the lowest pre-HRRP performance (Wasfy et al. 2017).

Hospital-Acquired Conditions Reduction Program The HACRP was effective beginning in fiscal year 2015. Hospitals are ranked on their total rate of preventable conditions in two categories: (1) claims-calculated patient safety indicators such as pressure ulcer and sepsis

(continued next page)

Redesigning Medicare's hospital quality and value payment programs

There is an opportunity to redesign Medicare's hospital quality payment programs as one hospital value incentive program (HVIP) that would be patient oriented, encourage coordination across providers and time, and promote change in the delivery system. Since current hospital quality programs are defined in statute, the Congress

rates and (2) hospital-reported health care–associated infections such as surgical site infections and catheter-associated urinary tract infections. The 25 percent of hospitals with the highest rates of preventable conditions (poorest performers) receive a 1 percent reduction to all inpatient payments. In 2017, the HACRP reduced payment to 742 hospitals, with penalties totaling roughly $370 million (Medicare Payment Advisory Commission 2017). Before the start of the HACRP, hospitals had been successful in reducing the number of hospital-acquired conditions (HACs). An Agency for Healthcare Research and Quality (AHRQ) study reported that, from 2010 to 2015, HACs per discharge declined by 21 percent, an estimated 125,000 fewer patients died in the hospital as a result of the reduction in HACs, and an estimated $28 billion in health care costs was saved (Agency for Healthcare Research and Quality 2016).

Hospital Value-based Purchasing Program The Hospital VBP Program was implemented in fiscal year 2013. As required by law, the program is budget neutral; that is, the total pool of withheld payments (currently equal to 2 percent of base inpatient DRG payments) must be redistributed to hospitals based on their performance on the VBP Program's quality measures. In 2018, the VBP Program increases payments to about 50 percent of IPPS hospitals and decreases payments to 42 percent of them. Hospitals earn back anywhere from 17 percent to 200 percent of their withheld payments. For roughly a third of these hospitals, the change in payments under the program was small, less than 0.25 percent of base payments.

The program uses a combination of measures from four quality domains to score hospitals on quality (the measures are also part of the IQRP): (1) 25 percent of the score is based on patient experience of care surveys; (2) 25 percent is based on patient safety, using a composite patient safety measure (AHRQ's patient safety indicator (PSI) 90) and data on six health care–associated infections; (3) 25 percent is based on efficiency, using a 30-day Medicare spending per beneficiary measure; and (4) 25 percent is based on clinical care, tied to 30-day mortality for three conditions—AMI, heart failure, and pneumonia).[14,15] The VBP Program gives a hospital credit for achievement (relative to other hospitals) and improvement (relative to its own baseline performance).

Public reporting of quality performance

Although not tied directly to payment, Medicare reports certain quality results to consumers and providers on CMS's Hospital Compare website. The website shows a hospital's results for given measure categories alongside the state and national averages for the measure. The displayed measures are from the IQRP, HRRP, HACRP, and VBP programs as well as results from hospital outpatient facilities (e.g., imaging efficiency). The measure categories include (1) survey of patient's experiences; (2) timely and effective care (i.e., cataract surgery care, heart attack care, emergency department care); (3) complications and deaths (e.g., health care–associated infections); (4) hospital returns; (5) use of medical imaging; and (6) payment and value of care (e.g., Medicare spending per beneficiary). The Hospital Compare website also presents a summary star rating (up to 5 stars) for the patient experience category and another star rating that combines individual clinical, patient experience, and efficiency measures from the VBP Program, HRRP, and the Hospital Compare website. The Commission has commented to CMS that the overall star rating system creates unneeded complexity in the Medicare program because it creates a new system of measures and scoring methodology for CMS to administer and for hospitals to track (Medicare Payment Advisory Commission 2016a). ∎

would need to create the new HVIP and eliminate the current programs in legislation. We believe that CMS has the authority to make some of our suggested changes to hospital quality payment without congressional action (e.g., improving public reporting.)[16] The HVIP is intended to replace quality programs that affect FFS hospital payment. However, in line with the Commission's principles, the HVIP measures and scoring methodology

should align across Medicare accountable entities and providers, including hospitals. MA plans, ACOs, and hospitals should be held accountable to a small set of population-based measures, scored against absolute thresholds, and have their payments adjusted through peer grouping. Medicare's use of the same set of measures and scoring framework across different populations could also promote multipayer alignment.

Design

The Medicare program should not pay hospitals and other providers for reporting quality measures, but should pay based on performance on these measures. Virtually all hospitals currently meet the IQRP reporting requirements and receive their full payment update, arguing for the need to retire the IQRP. The Congress could also consider removing payment incentives tied to Medicare quality reporting programs in other sectors where pay-for-performance programs have been implemented (e.g., skilled nursing facilities).

For simplicity, hospitals should have their payment adjusted based on performance on quality and cost measures in a single program instead of three separate programs. The HRRP and VBP programs should be combined into one HVIP. The HACRP, which scores patient safety measures such as infection rates, should also be retired as a hospital payment adjustment (see p. 194 for more discussion of patient safety).

Like the VBP Program, an HVIP would translate quality measure performance to payment and redistribute a budgeted amount to hospitals based on their performance. We would expect the new program to be budget neutral to the HRRP and HACRP, which, based on our analysis, reduce Medicare payment by 0.5 percent.

Public reporting of quality results can drive quality improvement by fostering competition among providers and allowing providers to better identify opportunities for improvement. We believe that CMS could incorporate an HVIP into the public reporting of quality results on Hospital Compare or other websites. CMS could report results as a consumer-friendly summary quality score (e.g., a star rating). For beneficiaries interested in more detailed quality results, CMS could also report all available patient experience measures (e.g., communication, cleanliness), some condition-specific outcomes (e.g., pneumonia readmissions, heart failure mortality), and HAC results.

Under an HVIP, the Medicare program would continue to provide hospitals with quality feedback reports to help them understand their performance on the claims-based measures. Reports could include benchmark and other comparative information so that hospitals could take action to improve their results. Even though an HVIP would score all-condition measures, CMS could consider providing hospitals with condition-specific results (e.g., acute myocardial infarction mortality) calculated by claims data for hospitals to use for their own quality improvement.

Measures

Based on our quality measurement principles, we propose an HVIP that would include four largely CMS-calculated or CMS-administered quality measures: mortality, readmissions, Medicare spending per beneficiary (MSPB), and patients' overall rating of the hospital.[17] These risk-adjusted measures are included in the existing hospital quality programs and thus are known to hospitals. (We envision that, as risk adjustment models evolve, they will be incorporated into the HVIP measures.) Providers could choose to use other granular quality measures to manage their own quality improvement, but these would not factor into Medicare payment.

Readmissions Hospital readmission, for any reason, is disruptive to patients and caregivers and costly to the health care system, and it puts patients at additional risk of hospital-acquired infections and complications. Readmissions are also a major source of patient and family stress and may contribute substantially to loss of functional ability, particularly in older patients. Measuring and adjusting payments based on a hospital's readmission rates holds the hospital accountable for ensuring that beneficiaries have the discharge information they need and encourages hospitals to coordinate with other providers. Since the implementation of the HRRP, hospitals have taken action and improved readmission rates. The readmission measure is also important to and understandable by the beneficiary and can be calculated through claims data.

In our HVIP model, we scored hospitals on their unplanned, risk-adjusted rates of readmissions within 30 days of discharge for all conditions using Medicare claims. Using an all-cause readmission measure (rather than the six conditions used in the HRRP) increases the number of observations and reduces random variation.

Our model also used three years of claims data (2014 through 2016) to increase the number of observations.

Mortality Mortality during or soon after a hospital stay (e.g., within 30 days) is an important outcome measure, and it encourages hospitals to coordinate with post-acute care providers. Like the readmission measure, this outcome measure can be determined with a high degree of accuracy through claims. As suggested with the readmissions measure, an all-condition mortality measure would hold hospitals more accountable than condition-specific measures. Our HVIP model used an all-condition, risk-adjusted measure of mortality during the hospital stay and 30 days after discharge, and we used three years of data (2014 to 2016) to increase the number of observations. (The measure excludes patients who are in hospice care before admission.)

Medicare spending per beneficiary MSPB is a claims-based value measure that we propose be included in an HVIP. This measure rewards efficient, effective hospital care, not volume of services, and reduces delivery system fragmentation. By pairing the spending measure with mortality and readmissions, hospitals have an incentive to maintain episode quality while reducing episode costs. The measure shows whether Medicare spends more, less, or about the same per Medicare patient treated at a specific hospital compared with how much Medicare spends on comparable patients nationally. Our model used the MSPB values CMS currently produces for the VBP Program, which are price-standardized, risk-adjusted episodes that include all Medicare Part A and Part B claims paid during the period from 3 days before an inpatient hospital admission through 30 days after discharge. The model used the MSPB values calculated with three years of data (2014 to 2016).

Patients' overall rating of the hospital The Hospital Consumer Assessment of Healthcare Providers and Systems® (HCAHPS®) is a national standardized survey instrument and data collection methodology for measuring patients' perspectives on their care during a recent hospital stay.[18] The survey allows Medicare, hospitals, beneficiaries, and others to make objective and meaningful comparisons of hospitals. Since 2006, CMS and hospitals have worked with third-party survey vendors to collect survey results from a random sample of each hospital's adult inpatient discharges. The survey results are used to calculate 10 core measures of patient experience: (1) communication with nurses,

(2) communication with doctors, (3) responsiveness of hospital staff, (4) communication about medicines, (5) cleanliness of hospital environment, (6) quietness of hospital environment, (7) discharge information, (8) care transition, (9) overall rating, and (10) whether the beneficiary would recommend the hospital to others. (Hospitals can add their own survey items to the core survey.) The HCAHPS measures are scored in the VBP Program; they are publicly reported on Hospital Compare and as part of the star rating system.

Based on the Commission's principles, a new HVIP ideally includes population-based patient experience measures. High-quality hospitals and physicians appear to focus not only on technical excellence but also on how patients perceive their care (Chatterjee et al. 2015). When patients have a better experience, they are more likely to adhere to treatments, return for follow-up appointments, and engage with the health care system by seeking appropriate care (Safran et al. 1998).

For simplicity, we modeled the HVIP using a single overall hospital rating measure (i.e., share of patients who gave their hospital a rating of 9 or 10 on a scale from 0 (lowest) to 10 (highest)) instead of a combination of the 10 HCAHPS measures. The overall hospital rating measure is strongly or moderately related to the other quality measures (e.g., communication with nurses correlation (r) = 0.64; care transition correlation (r) = 0.48), so by scoring a hospital's overall rating, we likely capture the other measures (Centers for Medicare & Medicaid Services 2017). Also, a hospital's performance on some of the other HCAHPS measures, such as discharge information and care transitions, would be detected in the readmissions, mortality, and MSPB measures. Alternatively, the HVIP could use a unique composite measure based on a subset of the HCAHPS measures that are meaningful to both beneficiaries and providers such as measures of communication with nurses, communication with doctors, responsiveness of staff, and discharge information.

To be scored on the overall hospital quality rating measure, hospitals would need to administer the entire core HCAHPS survey and would receive a score of zero for that measure if they did not. Hospitals could continue to monitor the other HCAHPS measures and use them to manage their own quality improvement. CMS could also continue to publicly report multiple HCAHPS measures on Hospital Compare.

Patient safety Our HVIP model adjusts a hospital's payment based on its performance on four measures that are part of the existing hospital quality payment programs. We also support a Medicare-influenced system to improve patient safety outside of an HVIP. But because of concerns with the accuracy of some patient safety data, we do not propose inclusion of patient safety measures in the HVIP model at this time. Under the HVIP, hospitals should continue to have incentives to improve patient safety because doing so could potentially affect performance on the four HVIP measures (e.g., readmissions due to hospital-acquired infections).

As part of the IQRP, HACRP, and VBP programs, hospitals are scored on five self-reported hospital care–acquired infection rates, such as catheter-associated urinary tract infections. Hospitals use their own claims and medical records to report their infection rates through the CDC's National Health Safety Network (NHSN).[19] The NHSN provides hospitals, states, and regions with comparative data needed to identify problem areas and measure local and national progress on prevention efforts. The monitoring and evaluation of infection rates through Medicare's programs and other national initiatives such as the Partnership for Patients have improved infection rates.

Over the years, there have been anecdotal reports of some hospitals' intentional misreporting of infection data—for example, clinicians ordering diagnostic tests in the absence of clinical symptoms to potentially identify infections present on admission so they are not considered hospital acquired (Centers for Disease Control and Prevention 2016a). The CDC and CMS have reported that there is no evidence such behaviors are widespread and have released guidance on the importance of adherence to the NHSN protocol, definitions, and criteria to ensure the reliability and comparability of the data. However, there are concerns that some hospitals are better than others at reporting infections and other patient safety issues (Calderwood et al. 2017). Also, even though there are specific definitions and criteria to capture the infection data, hospital infection control specialists have to make judgment calls about how to catalog infections, which makes part of the reporting subjective.

The IQRP, HACRP, and VBP programs also include a claims-based composite measure of 10 underlying patient safety indicators (PSIs), PSI 90, which signals potential in-hospital complications and adverse events and procedures, including pressure ulcers, iatrogenic pneumonia, postoperative sepsis, postoperative pulmonary embolism,

and accidental punctures or lacerations.[20] The use of the PSI 90 measure in pay-for-performance programs has been criticized for several reasons, including surveillance bias (e.g., hospitals with higher rates of postoperative blood clots were often the hospitals that were most vigilant in screening patients for them) and concerns about the accuracy of this measure in identifying meaningful unintentional cases of injury (Rajaram et al. 2015). AHRQ has recently updated the PSI 90 measure to address some of these concerns, and hospitals will begin to report on the revised measure this year. At this time, we do not propose to include the measure in the new payment program, but we will continue to monitor the measure's performance.

Hospital-acquired conditions are an important measure of patient safety, but since the only way currently to monitor a hospital's infection rate is through self-reported information, we propose that the current measures of infection rates not be part of a new HVIP. Rather, we suggest that hospitals be required as a Medicare condition of participation (COP) to report accurate infection rates to the NHSN and that hospitals continue to work with the CDC to monitor and evaluate opportunities to lower infection rates. (CMS could exempt small and rural hospitals that may not have sufficient patient numbers to warrant reporting to the NHSN.) This requirement can be built into the existing infection control COPs requiring hospitals to have a designated infection control officer, a hospital-wide quality assessment and performance improvement program, and training programs to address problems identified by the infection control officer (Centers for Medicare & Medicaid Services 2012). The Secretary should continue to publicly report infection rates (currently found on Hospital Compare) and investigate providers with high rates. Consistent with our principles, we also encourage CMS to support research and data collection to improve patient safety measures for potential inclusion in the HVIP.

Scoring methodology

Scoring under an HVIP should provide incentives for hospitals to improve the quality and efficiency of their care. To maintain the independence and importance of each of the four measures, our model treats each measure as an equally weighted, separate domain, consistent with the VBP Program methodology. Each of the 4 measures is worth 10 points for a total of 40 possible HVIP points. This model is illustrative; policymakers could give the components different weights based on the priorities of the Medicare program and its beneficiaries.[21]

TABLE
7–6

Illustration of point system to score performance on measures under our potential HVIP model

	Risk-adjusted readmissions rates (lower is better)	Risk-adjusted mortality rates (lower is better)	Relative Medicare spending per beneficiary (lower than 1 is better)	Patients' overall rating of hospital (higher is better)
0 points	20% or above	15% or above	1.16 or above	53% or below
2 points	18%	13%	1.09	60%
4 points	16%	11%	1.02	67%
6 points	14%	9%	0.95	73%
8 points	12%	7%	0.88	80%
10 points	10% or below	5% or below	0.82 or below	87% or above

Note: HVIP (hospital value incentive program). Each measure in the HVIP is continuously scored from 0 to 10 points, and only a subset of points is displayed here. Lower rates are better for readmissions, mortality, and Medicare spending per beneficiary (MSPB), and they receive more HVIP points. The MSPB value is based on the hospital's spending compared with the national mean. "Patients' overall rating of hospital" is the share of Hospital Consumer Assessment of Healthcare Providers and Systems® survey respondents who gave the hospital an overall rating of 9 or 10 on a 10-point scale.

Source: MedPAC analysis of Medicare fee-for-service hospital quality data, 2014–2016.

Converting measure performance to HVIP points (score)

One of the Commission's principles is that Medicare quality programs should reward providers based on clear, absolute, and prospectively set performance targets rather than score providers relative to one another. Prospective targets allow providers to know in advance what outcomes they must achieve to avoid penalties and achieve rewards; they also allow the industry as a whole to be rewarded if all providers improve. In addition, rewards should be distributed based on a continuous scale (i.e., without payment "cliffs"), so that hospitals with similar performance will receive similar financial rewards. In our example, hospitals earn points for their performance on quality metrics based on a continuous scale, starting at 0 points and gradually increasing to 10 points. The continuous scale stretches over almost the whole distribution of performance, giving even top-performing hospitals an incentive to continue to improve.

In our HVIP model, each measure has a continuous performance-to-points scale based on the 2nd percentile of hospital performance (0 points) to the 98th percentile of hospital performance (10 points), which is based on the hospitals in our data set. This scale—from the 2nd percentile to 98th percentile—is meant to represent empirically derived scores that available evidence suggests can be achieved by an optimally performing hospital (Safran et al. 2007). Although scoring is continuous, hospitals would know in advance what performance

targets (or "gates") they needed to reach to achieve a certain point level for each measure. Table 7-6 presents a subset of the scale of points associated with performance targets in our HVIP model.

Following is an example of converting measure performance to points using the continuous performance-to-points scale highlighted in Table 7-6: Hospital A has a risk-adjusted readmissions rate of 15 percent (earns 5 points), risk-adjusted mortality rate of 7 percent (earns 8 points), Medicare spending per beneficiary value of 0.96 (earns 5.9 points), and overall patient experience rating of 79 percent (earns 7.8 points). Hospital A receives a total of 26.7 of 40 possible HVIP points.

Each hospital's total quality performance score, which would be used to determine its HVIP payment adjustment, would have a maximum of 40 points. In our HVIP model, each hospital has a total number of points based on its performance against our continuous performance-to-points scale (Table 7-6). The 3,021 hospitals included in our sample had a nearly normal distribution of total quality performance scores under our HVIP model (Figure 7-2, p. 196).[22]

In our HVIP model, the average total HVIP score point total for all hospitals was 22.9 points (Table 7-7, p. 197). On average, mortality contributed 7 of those points because more hospitals perform better on this measure

FIGURE
7–2

**Hospitals have a nearly normal distribution of
total quality performance under the potential HVIP**

Total HVIP points (out of 40 points)
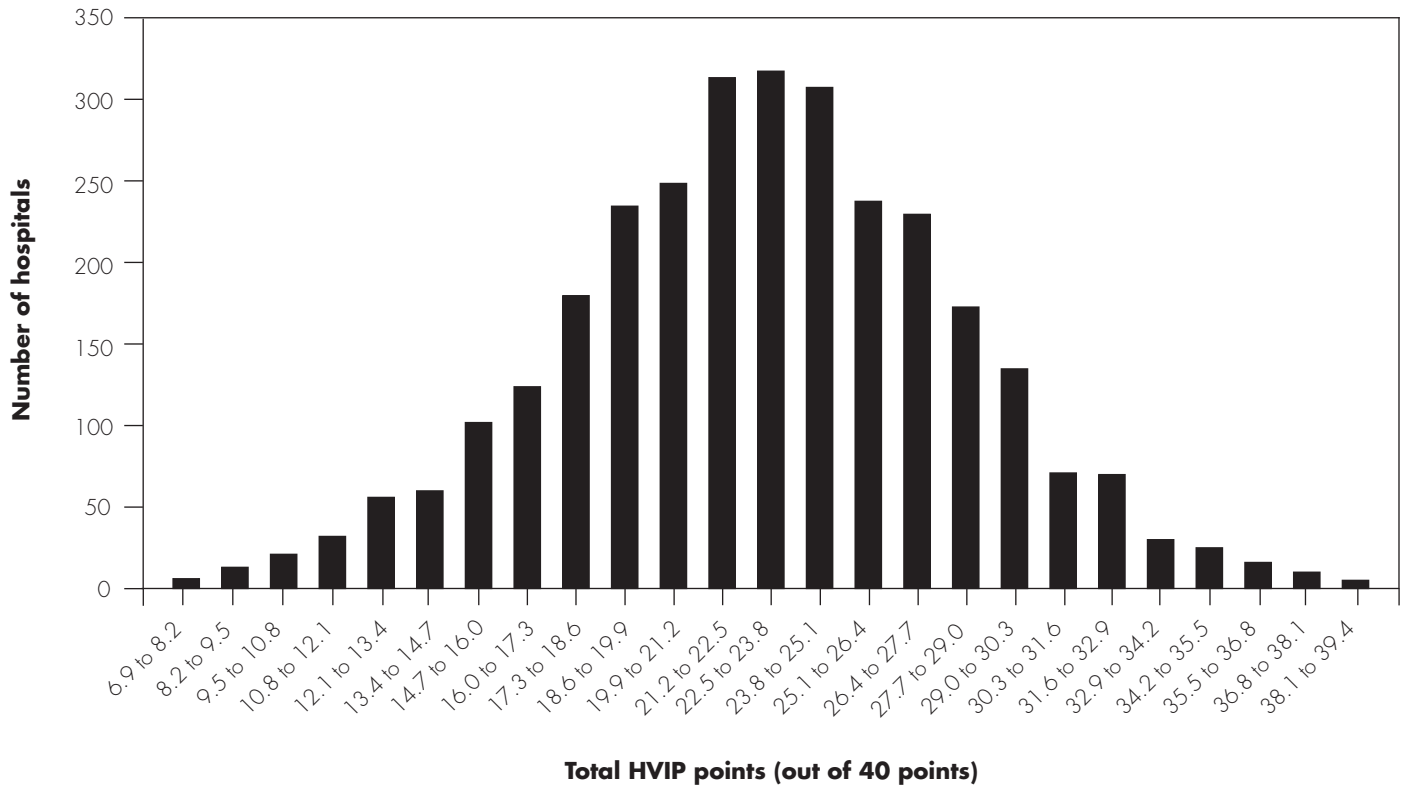
Note: HVIP (hospital value incentive program). Hospitals receive 0 to 40 total HVIP points based on their performance on four equally weighted measures (readmissions, mortality, Medicare spending per beneficiary, and patients' overall rating of hospital). There are 3,021 hospitals included in our HVIP model.

Source: MedPAC analysis of Medicare fee-for-service hospital quality data, 2014–2016.

compared with readmissions, MSPB, and overall patient experience rating, which each contributed about 5 points to the total score. In addition, there were some differences in total HVIP scores based on hospital characteristics. For example, in our model, major teaching hospitals had a lower average total HVIP score compared with nonteaching hospitals (21.2 points compared with 23.2 points, respectively). This difference is partially because major teaching hospitals have worse readmission rates and therefore fewer points in that domain of the HVIP scoring model (3.8 points for teaching hospitals compared with 5.5 points for nonteaching hospitals).

Converting HVIP points to payment adjustments using peer grouping In measuring providers' performance on quality measures, the Commission contends that Medicare should take into account, as necessary, differences in

providers' populations, including social risk factors. However, adjusting measure results for social risk factors can mask disparities in clinical performance, so Medicare should adjust performance payments through peer grouping rather than through performance score adjustments. (In peer grouping, each provider is compared with its "peers"—defined as providers with a similar patient mix.)[23] The Commission also believes that Medicare should target technical assistance resources to low-performing providers and should support research and data collection to reduce measurement bias, including, for example, the effects of social risk factors.

Based on these principles, our HVIP model distributes quality-based payments to hospitals classified in 10 peer groups. Each peer group has about the same number of hospitals (in our model, about 300 hospitals), and

TABLE
7–7

TABLE 7–7

Illustrative total HVIP points by hospital characteristics

Hospital group	Number of hospitals	Total HVIP points (score)	Readmissions points	Mortality score	MSPB score	Patients' overall rating of hospital
			Average:			
All hospitals	3,021	22.9	5.3	7.0	5.1	5.5
Hospital size						
Large urban	1,209	22.2	4.8	7.6	4.5	5.3
Other urban	1,065	23.8	5.7	7.2	5.2	5.7
Rural	747	22.9	5.5	5.9	6.1	5.4
Teaching status						
Major teaching	300	21.2	3.8	7.8	4.7	4.9
Other teaching	764	22.9	5.2	7.6	4.9	5.3
Nonteaching	1,957	23.2	5.5	6.7	5.3	5.6
Ownership						
Nonprofit	1,826	23.9	5.5	6.8	5.3	5.7
For profit	754	21.1	4.8	7.4	4.4	5.1
Government	441	22.1	5.2	6.1	5.5	5.3
DSH						
No DSH	410	25.8	6.1	7.2	5.5	7.1
Moderate to low DSH	1,897	23.2	5.5	6.9	5.3	5.5
High DSH	665	20.3	4.2	7.3	4.5	4.3

Note: HVIP (hospital value incentive program), MSPB (Medicare spending per beneficiary), DSH (disproportionate share). Hospitals receive up to a total of 40 points based on their performance on four equally weighted measures (up to 10 points each): risk-adjusted, unplanned readmissions; risk-adjusted 30-day postdischarge mortality; MSPB; and patients' overall rating of hospital. "Patients' overall rating of hospital" is the share of Hospital Consumer Assessment of Healthcare Providers and Systems® survey respondents that gave the hospital an overall rating of 9 or 10 on 10-point scale. "High DSH" hospitals have higher proportions of low-income patients compared to "no DSH" hospitals. There are 49 hospitals with unknown DSH status.

Source: MedPAC analysis of Medicare fee-for-service hospital quality data, 2014–2016.

hospitals are assigned to peer groups based on their share of Medicare patients who are fully dual-eligible beneficiaries—that is, who also fully qualify for Medicaid, which can be a proxy for low income. (In fiscal 2019, the HRRP will use five peer groups based on the hospital's share of Medicare patients who are fully dual-eligible beneficiaries.) Since our HVIP model is designed to be budget neutral, each peer group has, in essence, a budget based on a 2 percent payment withhold from each of the peer group's hospitals. This budget is redistributed to the peer group's hospitals based on their HVIP points. The 2 percent withhold is the same as the withhold in the existing VBP Program, but policymakers could raise or lower that amount.

In our HVIP model, we followed five steps to convert performance points to payment adjustments using currently available hospital quality and payment data. (See text box, pp. 202–203, describing the process used in our HVIP model to convert each hospital's HVIP points to a quality-based payment adjustment.) Overall, we found that it was feasible to compute incentive payments that support the Commission's HVIP's goals.

After scoring each hospital on the same continuous performance-to-points scale, we divided the 3,021 hospitals in our HVIP sample into 10 equal-sized peer groups based on the hospitals' shares of fully dual-eligible Medicare patients (text box Steps 1 and 2, p. 202). The

Peer group	Average:		Peer group budget based on 2 percent withhold of hospitals' base IPPS payments (in millions)	Percentage adjustment to base IPPS payments per HVIP point
	Share of fully dual-eligible beneficiaries	Total HVIP points		
1 (lowest share of fully dual-eligible beneficiaries)	6.5%	26.4	$211.8	0.08%
2	10.8	24.5	228.2	0.08
3	13.1	23.9	274.9	0.08
4	15.2	23.7	227.8	0.08
5	17.2	23.7	208.6	0.08
6	19.3	22.6	216.4	0.09
7	22.1	22.6	165.8	0.09
8	25.3	22.3	169.5	0.09
9	30.5	21.2	179.5	0.09
10 (highest share of fully dual-eligible beneficiaries)	48.3	18.3	148.3	0.10

Note: HVIP (hospital value incentive program), IPPS (inpatient prospective payment system). There are about 300 hospitals in each of the 10 hospital peer groups. Peer groups are assigned based on the share of the hospital's Medicare patients who are fully dual eligible for Medicare and Medicaid for a majority of the year. Fully dual-eligible beneficiaries qualify for a full range of Medicaid benefits.

Source: MedPAC analysis of Medicare fee-for-service hospital quality data, 2014–2016.

average share of these beneficiaries per hospital in each peer group ranged from less than 7 percent (Peer Group 1) to about 48 percent (Peer Group 10) (Table 7-8). The average total HVIP points hospitals received in each peer group ranged from 26.4 (Peer Group 1) to 18.3 (Peer Group 10). Peer Group 10 had fewer total HVIP points mainly because of higher average readmissions and lower overall patient ratings compared with Peer Group 1 hospitals. Although, on average, Peer Group 10's point total was lower, some hospitals in the peer group were high performers and received more HVIP points than the average hospital.

For each peer group, we calculated a budget of expected HVIP payments to the group's hospitals based on a 2 percent withhold of base inpatient prospective payment system (IPPS) payments from each of the group's hospitals (text box Step 3, p. 202). Under our model, a total of $2.03 billion is distributed to hospitals based on their HVIP points. The budget for each peer group ranges from about $275 million (Peer Group 3) to about $148 million (Peer Group 10) (Table 7-8). Inherent in the peer group budgets are the number of discharges for the peer group's hospitals, so the budget is smaller for those peer groups that have hospitals with fewer discharges and thus

smaller IPPS base payments to be used in the withhold calculation.

For each peer group, we also calculated the percentage adjustment to payments per point, which converts total HVIP points to dollars and results in spending the 2 percent withhold for each group (text box Step 4, p. 202). The percentage adjustments to payments per point range from 0.08 percent (Peer Group 1) to 0.10 percent (Peer Group 10) (Table 7-8). In other words, high-performing hospitals in Peer Group 10 have the potential to earn a slightly higher payment adjustment per performance point compared with the other groups because the percentage adjustment to payments per point for Peer Group 10 is higher than the other groups.

We calculated each hospital's HVIP-based payment adjustment using its total HVIP points and its peer group's conversion factor for points-to-payment adjustment (text box Step 5, p. 202). In our HVIP model, small differences exist between the peer groups' ranges of payment adjustments. In general, a hospital's payment adjustment could range from –1.4 percent to 1.6 percent based on the hospital's base IPPS payment after accounting for their 2 percent withhold (Table 7-9). (By design, no hospital can

198 Applying the Commission's principles for measuring quality: Population-based measures and hospital quality incentives

MEDPAC

TABLE
7–9

Illustrative HVIP payment adjustments by hospital peer groups

Peer group	Average withhold of total base IPPS payments	Range of HVIP payment adjustments	
		After withhold	Relative to withhold
1 (lowest share of fully dual-eligible beneficiaries)	2%	–1.1% to + 1.1%	44% to 156%
2	2	–1.1 to + 1.1	47 to 155
3	2	–1.2 to + 1.0	38 to 149
4	2	–1.1 to + 0.9	44 to 146
5	2	–1.2 to + 1.1	40 to 152
6	2	–1.1 to + 1.0	45 to 152
7	2	–1.3 to + 1.1	37 to 154
8	2	–1.4 to + 1.3	31 to 163
9	2	–1.3 to + 1.2	37 to 158
10 (highest share of fully dual-eligible beneficiaries)	2	–1.3 to + 1.6	37 to 180

Note: HVIP (hospital value incentive program), IPPS (inpatient prospective payment system). There are about 300 hospitals in each of the 10 hospital peer groups. Peer groups are assigned based on the share of the hospital's Medicare patients who are fully dual eligible for Medicare and Medicaid for a majority of the year. Fully dual-eligible beneficiaries qualify for a full range of Medicaid benefits. The average HVIP adjustments after the withhold is zero by design.

Source: MedPAC analysis of Medicare fee-for-service hospital quality data, 2014–2016.

lose more than its 2 percent withhold.) Thus, hospitals can recover between 31 percent and 180 percent of their 2 percent withhold.

Under our model, about half of the hospitals (1,510) would receive a penalty and about half (1,511 hospitals) would receive a reward. About 11 percent of hospitals (367) would receive a reward more than 1.5 times the withhold. About 12 percent (365) would receive a penalty of less than one-half of the withhold.

Comparison of HVIP model to existing hospital quality programs

To understand differences between hospital performance in the existing programs and our HVIP model, we assigned hospitals to quartiles based on their total performance in the existing programs and then quartiles based on their performance under the HVIP model. About a quarter of hospitals were in the same performance quartile under the existing programs and the HVIP model. Three-quarters of hospitals were in the same or within one performance quartile under the existing program and the HVIP model. At the extremes, 2 percent (61 hospitals) were poor performers in the existing programs but were

top performers in the HVIP model. About 1 percent (34 hospitals) were top performers in the existing programs but were poor performers in the HVIP model. The HACRP appeared to play a role in this trend (i.e., some hospitals were poor performers in the existing programs because they received a HAC penalty but did well under the HVIP model.) This supports our concerns with potential misreporting of hospital infection data in a program that uses a tournament model rather than fixed targets.

Effect of peer grouping on reducing disparities among hospitals

Our HVIP model uses a small set of measures, a continuous performance-to-points scale, and converts those points to payment adjustments relative to groups of hospitals that serve similar shares of fully dual-eligible populations (hospital peer groups). Since one goal of an HVIP is to adjust payments to account for differences in social risk factors, we examined how hospitals serving large shares of low-income patients perform.[24] Figure 7-3 (p. 200) compares the existing quality payment program adjustments with the HVIP model's payment adjustments by peer group. All the HVIP adjustments are zero relative to the average within each peer group since

FIGURE
7–3

Compared with existing quality payment programs, the potential HVIP makes payment adjustments more equitable for hospitals grouped by share of fully dual-eligible beneficiaries
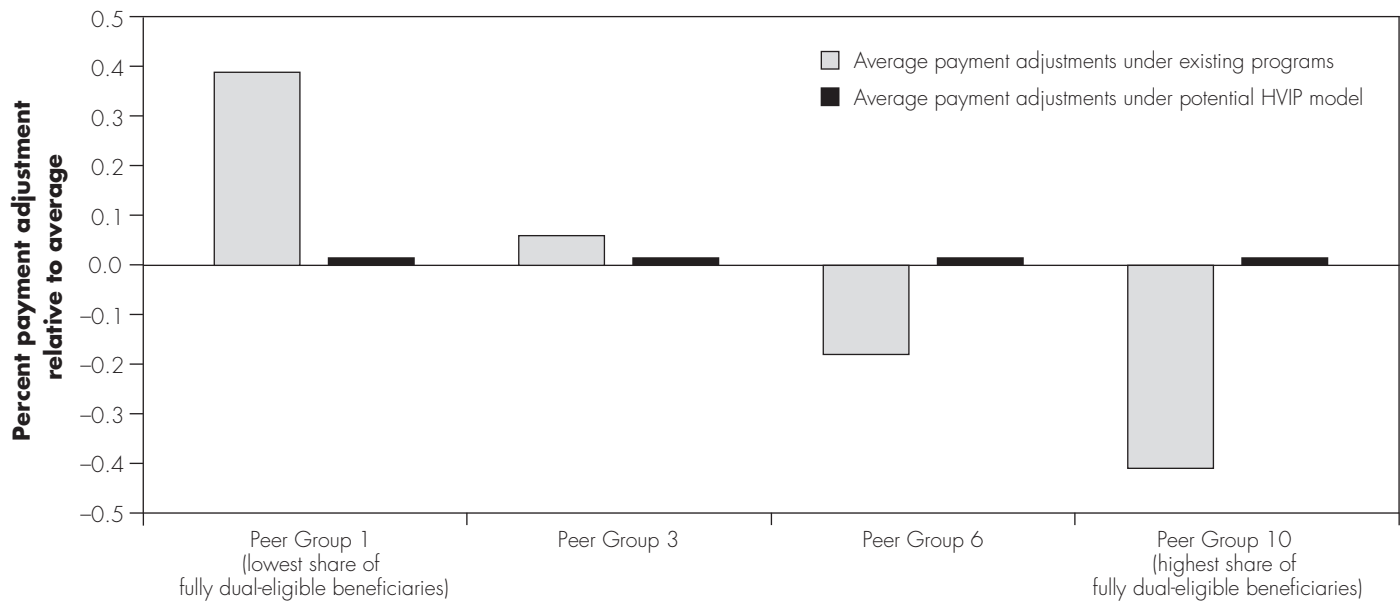
Note: HVIP (hospital value incentive program). The existing quality programs include the Hospital Readmissions Reduction Program (HRRP), Hospital-Acquired Condition Reduction Program (HACRP), and Hospital Value-based Purchasing (VBP) Program. The HRRP and HACRP are penalties, and the VBP Program is budget neutral. To make the existing programs and HVIP comparable, we included a budget-neutrality adjustment in the existing programs' adjustment. The budget-neutrality adjustment is the overall existing program adjustment divided by overall base payments (0.93 percent). The average HVIP adjustment is the sum of each hospital's HVIP adjustment after the withhold divided by the sum of each hospital's base payment. The HVIP is budget neutral. Peer groups are assigned based on the share of the hospital's Medicare patients who are fully dual eligible for Medicare and Medicaid for a majority of the year. Fully dual-eligible beneficiaries qualify for the full range of Medicaid benefits.

Source: MedPAC analysis of Medicare fee-for-service hospital quality data, 2014–2016.

the adjustments are budget neutral within each peer group. Under the existing programs, Peer Group 1 (lowest share of fully dual-eligible beneficiaries) hospitals receive a 0.39 percent positive adjustment while Peer Group 10 (highest share of fully dual-eligible beneficiaries) hospitals receive a –0.41 percent adjustment. Thus, compared with the existing quality payment programs, the HVIP approach makes payment adjustments among hospitals that serve different populations more equitable.

We can also see this effect in Figure 7-4, which compares existing and HVIP model payment adjustments for different groups of hospitals according to their disproportionate share (DSH) hospital status (which can also be considered a proxy for low-income status). Under the existing quality programs, non-DSH hospitals receive, on average, a 0.42 percent positive adjustment; under an

HVIP program, they would receive a smaller positive adjustment—on average, 0.06 percent. Under the existing programs, the high–DSH hospitals receive, on average, a –0.22 percent adjustment; under an HVIP program, that adjustment would rise to an average of –0.04 percent.

Conclusion

A single quality payment program for hospitals, such as our HVIP model, would be simpler to administer and would produce more equitable results compared with the existing quality payment programs. The HVIP, as a single program, would eliminate the complexity of overlapping program requirements, focus on outcomes, and promote the coordination of care. It would also align with the Commission's principles for quality measurement, in particular, by setting absolute value targets and using

FIGURE
7–4

Compared with existing quality payment programs, the potential HVIP makes payment adjustments more equitable for hospitals grouped by DSH status
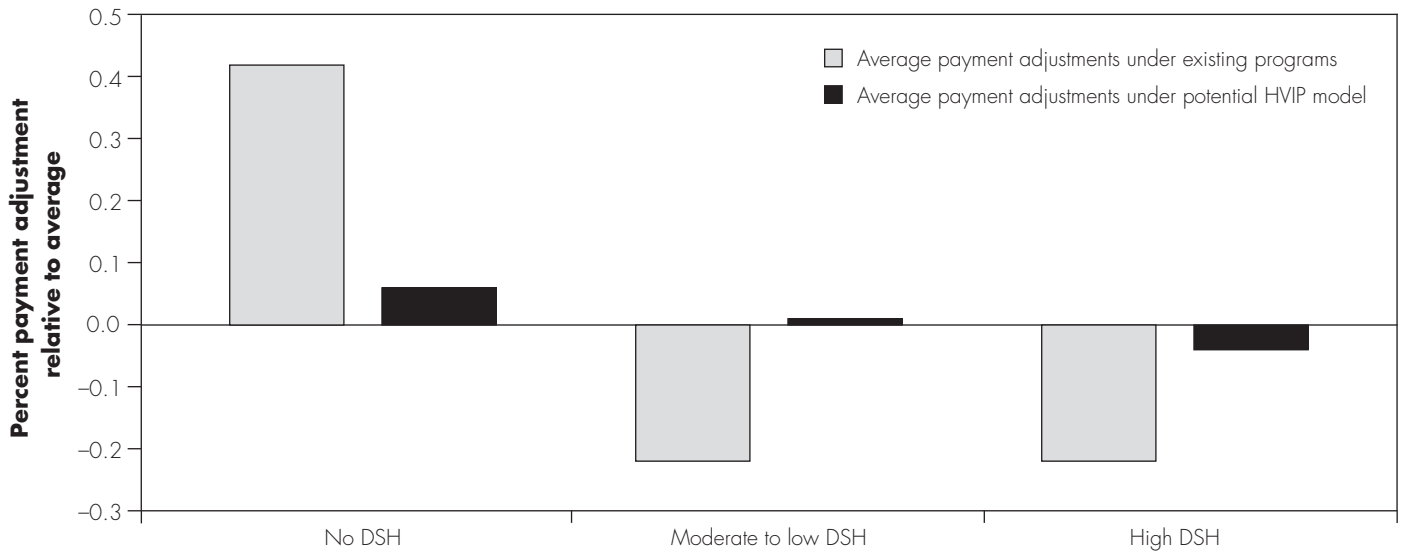
Note: HVIP (hospital value incentive program), DSH (disproportionate share hospital). The existing quality programs include the Hospital Readmissions Reduction Program (HRRP), Hospital-Acquired Condition Reduction Program (HACRP), and Hospital Value-based Purchasing (VBP) Program. The HRRP and HACRP are penalties, and the VBP Program is budget neutral. To make the existing programs and HVIP comparable, we included a budget-neutrality adjustment in the existing programs' adjustment. The budget-neutrality adjustment is the overall existing program adjustment divided by overall base payments (0.93 percent). The average HVIP adjustment is the sum of each hospital's HVIP adjustment after the withhold divided by the sum of each hospital's base payment. The HVIP is budget neutral.

Source: MedPAC analysis of Medicare fee-for-service hospital quality data, 2014–2016.

peer grouping to account for differences in provider populations. Under peer grouping in our HVIP model, differences in payment adjustments were reduced among providers serving populations of varying social risk factors.

Over the next year, the Commission plans to continue to refine a design for an HVIP consistent with our principles for quality measurement. Some topics the Commission will further explore include weighting of measures, withhold values, patient experience measures, and patient safety measures. ∎

Steps to convert hospital value incentive program points to payment adjustments using peer grouping

Our hospital value incentive program (HVIP) model distributes quality-based payments to hospitals classified in 10 peer groups. Each peer group has about the same number of hospitals, and hospitals are assigned to peer groups based on their share of Medicare patients who are fully dual-eligible beneficiaries—that is, who also fully qualify for Medicaid, which can be a proxy for low income. Since our HVIP model is designed to be budget neutral, each peer group has, in essence, a budget based on a 2 percent payment withhold from each of the peer group's hospitals. This budget is redistributed to the peer group's hospitals based on their quality performance.

We followed five steps to covert each hospital's quality measure performance to a payment adjustment that provides rewards or penalties.

Step 1: Convert each hospital's performance on quality measures to total HVIP points based on a continuous performance-to-points scale. (Every hospital is scored on the same scale.)

Step 2: Divide hospitals into 10 equal-sized peer groups based on the hospital population's share of fully dual-eligible patients.

Step 3: For each peer group, create a budget of expected HVIP payments to hospitals, based on a 2 percent withhold from each of the hospitals in the peer group (e.g., 2 percent of each hospital's base inpatient prospective payment system (IPPS) payments).

Step 4: For each peer group, calculate the percentage adjustment to payment per HVIP point, which converts total HVIP points to dollars and results in spending the group's budget defined in Step 3.

Percentage adjustment to payments per point = HVIP budget for peer group / (sum (each hospital's base IPPS payments × hospital's total HVIP points))

Step 5: Compute each hospital's adjustment for the coming year based on past performance and their peer group's percentage adjustment to payment per HVIP point.

Hospital's HVIP-based adjustment = percentage adjustment to payments per point × hospital's total HVIP points

Multiply the hospital's HVIP-based adjustment by the hospital's withhold of IPPS payments to yield the payment adjustment in dollars.

(continued next page)

Steps to convert hospital value incentive program points to payment adjustments using peer grouping (cont.)

Table 7-10 below describes an example of converting HVIP points to payment adjustments using peer grouping. First, we convert each hospital's quality measure performance to total HVIP points based on the continuous performance-to-points scale (Step 1). As seen at the top of the table, Hospital 1 has higher total HVIP performance with 40 points compared with Hospital 2's 30 points. We assume two hospitals were assigned to a peer group because of a similar share of fully dual-eligible beneficiaries (Step 2). We withhold 2 percent of each of the hospital's total base IPPS payments (Step 3). Since Hospital 1 has fewer discharges, its 2 percent withhold is less than Hospital 2's withhold. As shown in the middle of the table, the total HVIP bonus pool to be redistributed

for the peer group is a sum of the two hospital's withholds (or $1.3 million). We then calculate the percentage adjustment to payments per point for the peer group, which converts total HVIP points to dollars and results in spending the entire $1.3 million budget (Step 4). For every HVIP point that a hospital in the peer group earns, it can receive a 0.065 percent payment adjustment per point. Based on the hospital's HVIP performance and the peer group's percentage adjustment to payments per point, Hospital 1 will earn a payment adjustment of 2.6 percent, which is equal to $130,000 (or a reward of $30,000 greater than the hospital's withhold) (Step 5). Because Hospital 2 had lower HVIP points, it will have a $30,000 penalty. ■

TABLE 7–10

Example of converting HVIP points to payment adjustments for a peer group's hospitals

	Hospital 1 (500 discharges)	Hospital 2 (5,000 discharges)
HVIP points (Step 1)	40	30
Total base IPPS payments	$5,000,000	$60,000,000
2 percent withhold of IPPS payments	$100,000	$1,200,000
Total HVIP budget for peer group (Step 3)	$1,300,000	
Percentage adjustment to payments per point (Step 4)	0.065% adjustment per point	
Hospital HVIP-based adjustment (Step 5)	2.60% ($130,000)	1.95% ($1,170,000)
Reward or penalty relative to 2 percent withhold	+0.60% (+$30,000)	–0.05% (–$30,000)

Note: HVIP (hospital value incentive program), IPPS (inpatient prospective payment system). This example assumes the peer group has two hospitals (Step 2).

Endnotes

1   For clarity and consistency with the Commission's past work, we use the term *potentially preventable admissions* throughout the chapter. The literature and industry also refer to the measure concept as *avoidable hospitalizations, ambulatory care–sensitive condition hospitalizations,* and *hospitalizations for potentially preventable complications.*

2   HEDIS is a registered trademark of NCQA. The HEDIS potentially preventable admissions measure is called "hospitalizations for potentially preventable complications."

3   CMS has proposed to retain this measure as a 2019 MA Plan Finder display page measure. The agency has also signaled its intent to move the measure to the star rating program in 2022 (Centers for Medicare & Medicaid Services 2018a).

4   The expected-discharges value is the predicted number of hospitalizations based on the age, sex, and comorbidities (i.e., hierarchical condition categories (HCCs)) of the eligible population of beneficiaries.

5   The measure uses discharges rather than admissions because patients who die in the hospital are not included in the measure. For consistency, we use the term *potentially preventable admissions*.

6   Eight percent represents 740,000 potentially preventable admissions out of 9.5 million admissions.

7   The Commission has previously referred to this measure as "healthy days at home." The measure's new name does not presume that beneficiaries are healthy just because they are at home and is more explicitly inclusive of beneficiaries who may be living in long-term care facilities.

8   For example, the Endovascular Treatment for Small Core and Anterior Circulation Proximal Occlusion with Emphasis on Minimizing CT to Recanalization Times (ESCAPE) study uses a days alive and out of the hospital measure during the six months after the randomized use of pulmonary artery catheters for patients with congestive heart failure.

9   Because our goal is to calculate market-specific estimates of HCDs and ultimately compare payment models across and within market areas, we used a fixed-effect model that includes an indicator variable for each of the markets in the regression model to better estimate the age, sex, and HCC covariates.

10 *Market areas* refers to the over 1,200 MedPAC-defined market areas used in the PPA analysis.

11 The HCD measure includes beneficiaries ages 65 years and older, while the PPA measure was specified for beneficiaries ages 67 years and older. The PPA measure focuses on admissions tied to five chronic conditions. For the HCD calculations, chronic conditions are identified from a set of 15 (acute myocardial infarction/ischemic heart disease, CHF, specified heart arrhythmias, dementia, hematologic disease, lung disease, psychiatric disease, chronic kidney disease, endocrine disease, vascular disease, neuromuscular disease, diabetes, cancer, liver disease, stroke). The conditions were chosen based on the combination of high prevalence and mortality as well as associated health care spending.

12 One possible explanation for the increase in mortality days in 2015 is the very severe flu season from October 2014 to March 2015. Beneficiaries who died in the January to March portion of the 2014 to 2015 flu season would have fewer HCDs because they had more mortality days subtracted from the 365 calendar days of 2015. Beneficiaries who died in the October to December portion of the 2014 to 2015 flu season would have more HCDs because they had fewer mortality days subtracted from the 365 calendar days of 2014.

13 The Commission recommended a readmissions reduction program in our 2008 report to the Congress (Medicare Payment Advisory Commission 2008). Our June 2018 report to the Congress also includes a study mandated by the 21st Century Cures Act of 2016 that examines whether changes in readmission rates under the HRRP are related to any changes in outpatient and emergency services furnished.

14 The PSI 90 measure is a composite of eight patient safety measures: PSI 03 (pressure ulcer); PSI 06 (iatrogenic pneumothorax); PSI 07 (central venous catheter–related bloodstream infections); PSI 08 (postoperative hip fracture); PSI 12 (perioperative pulmonary embolism or deep vein thrombosis); PSI 13 (postoperative sepsis); PSI 14 (postoperative wound dehiscence); and PSI 15 (accidental puncture or laceration).

15 In 2018, 2 process-of-care measures were dropped from the VBP Program, and the 1 remaining process-of-care measure, PC–01 (elective delivery before 39 weeks), was moved to the patient safety domain; this measure's weight increased from 20 percent to 25 percent. CMS has proposed removing the PC–01 measure from the VBP Program (Centers for Medicare and Medicaid Services 2018b).

16 The IQRP was mandated by the Medicare Prescription Drug, Improvement, and Modernization Act of 2003 and updated by the Patient Protection and Affordable Care Act of 2010 (PPACA). The HRRP, VBP Program, and HACRP are mandated in PPACA.

17 CMS calculates claims-based mortality, readmissions, and MSPB measures. CMS oversees the Hospital Consumer Assessment of Healthcare Providers and Systems® patient experience survey (including certifying survey vendors and developing standardized data collection and sampling protocols). Hospitals work with a survey vendor or follow the standardized protocols themselves to collect and report the core and supplemental experience data from their patients.

18 CAHPS® is a registered trademark of AHRQ, a U.S. government agency.

19 NHSN is operated by the Centers for Disease Control and Prevention. It is the nation's most widely used health care–associated infection tracking system. Acute care hospitals, long-term care hospitals, dialysis facilities, inpatient rehabilitation facilities, and other facility types report data to NHSN.

20 This PSI 90 composite measure was created by AHRQ to help hospitals measure adverse events and address their own quality improvement efforts.

21 In some current Medicare quality payment programs, CMS uses differential measure weighting to prioritize outcome measures over process measures. For example, the MA star rating program assigns outcome and intermediate outcome measures a weight of 3, patient experience and access measures a weight of 1.5, and process measures a weight of 1.

22 We included only hospitals paid through the inpatient prospective payment system. Because we wanted to model the scoring of all four measures, we did not include hospitals with no publicly reported HCAHPS data or MSPB data (from CMS) or risk-adjusted mortality or readmissions value of 0 or missing. A policy question is how to score missing values—for example, when a hospital's population is too small for HCAHPS. Another policy question is whether and how to include critical access hospitals, which may have numbers too small for valid measurement.

23 Based on suggestions from the Commission and the recent requirement legislated in the 21st Century Cures Act of 2016, CMS is implementing a peer-group scoring model, using five peer groups, in the HRRP. Others have tested and found that the peer-grouping approach adequately accounts for differences among providers serving populations with social risk factors (Office of the Assistant Secretary for Planning and Evaluation 2016, Samson et al. 2018).

24 We compared the amount of quality payment adjustments in existing programs with the HVIP model payment adjustments by hospital characteristics (e.g., size, teaching status) (see Table 7-A1 in online Appendix 7-A, available at http://www.medpac.gov). To make the existing programs and HVIP comparable, we included a budget-neutrality adjustment for the existing program adjustment calculation.

References

Agency for Healthcare Research and Quality. 2016. *National scorecard on rates of hospital-acquired conditions 2010 to 2015: Interim data from national efforts to make health care safer*. Rockville, MD: AHRQ. http://www.ahrq.gov/professionals/quality-patient-safety/pfp/2015-interim.html.

Barnett, M. L., D. C. Grabowski, and A. Mehrotra. 2017. Home-to-home time: Measuring what matters to patients and payers. *New England Journal of Medicine* 377, no. 1 (July 6): 4–6.

Burns, M., and J. Mullahy. 2016. *Healthy-time measures of health outcomes and healthcare quality*. NBER working paper no. 22562. Cambridge, MA: National Bureau of Economic Research.

Calderwood, M. S., S. S. Huang, V. Keller, et al. 2017. Variable case detection and many unreported cases of surgical-site infection following colon surgery and abdominal hysterectomy in a statewide validation. *Infection Control and Hospital Epidemiology* 38, no. 9 (September): 1091–1097.

Centers for Disease Control and Prevention. 2016a. Adherence to the Centers for Disease Control and Prevention's (CDC's) infection definitions and criteria is needed to ensure accuracy, completeness, and comparability of infection information. https://www.cdc.gov/nhsn/cms/cms-reporting.html.

Centers for Disease Control and Prevention. 2016b. *National and state healthcare-associated infections progress report*. Atlanta, GA: CDC.

Centers for Medicare & Medicaid Services, Department of Health and Human Services. 2018a. Advance notice of methodological changes for calendar year (CY) 2019 for Medicare Advantage capitation rates, Part C and Part D payment policies and 2019 draft call letter. February 1. https://www.cms.gov/medicare/health-plans/medicareadvtgspecratestats/downloads/advance2018.pdf.

Centers for Medicare & Medicaid Services, Department of Health and Human Services. 2018b. Fiscal year (FY) 2019 Medicare hospital inpatient prospective payment system (IPPS) and long term acute care hospital (LTCH) prospective payment system proposed rule, and request for information. https://www.cms.gov/Newsroom/MediaReleaseDatabase/Fact-sheets/2018-Fact-sheets-items/2018-04-24.html.

Centers for Medicare & Medicaid Services, Department of Health and Human Services. 2017. HCAHPS patient-level correlations table. http://www.hcahpsonline.org/globalassets/hcahps/summary-analyses/patient-level/july-2015--june-2016-discharges.pdf.

Centers for Medicare & Medicaid Services, Department of Health and Human Services. 2012. Medicare program; reform of hospital and critical access hospital conditions of participation. Final rule. *Federal Register* 77, no. 95 (May 16): 29034–29076.

Chatterjee, P., T. C. Tsai, and A. K. Jha. 2015. Delivering value by focusing on patient experience. *American Journal of Managed Care* 21, no. 10 (October): 735–737.

Covinsky, K. E., E. Pierluissi, and C. B. Johnston. 2011. Hospitalization-associated disability: "She was probably able to ambulate, but I'm not sure." *Journal of the American Medical Association* 306, no. 16 (October 26): 1782–1793.

Davies, S. M., K. M. McDonald, E. Schmidt, et al. 2009. *Expanding use of the Prevention Quality Indicators: Report of Clinical Expert Review Panel*. Report prepared for the Agency for Healthcare Research and Quality. Rockville, MD: AHRQ.

Freund, T., S. M. Campbell, S. Geissler, et al. 2013. Strategies for reducing potentially avoidable hospitalizations for ambulatory care-sensitive conditions. *Annals of Family Medicine* 11, no. 4 (July–August): 363–370.

Gao, J., E. Moran, Y. F. Li, et al. 2014. Predicting potentially avoidable hospitalizations. *Medical Care* 52, no. 2 (February): 164–171.

Medicare Payment Advisory Commission. 2017. *Report to the Congress: Medicare payment policy*. Washington, DC: MedPAC.

Medicare Payment Advisory Commission. 2016a. Letter to the Acting Administrator of the Centers for Medicare and Medicaid Services regarding the hospital star rating system. September 22.

Medicare Payment Advisory Commission. 2016b. *Report to the Congress: Medicare payment policy*. Washington, DC: MedPAC.

Medicare Payment Advisory Commission. 2015a. *Report to the Congress: Medicare and the health care delivery system*. Washington, DC: MedPAC.

Medicare Payment Advisory Commission. 2015b. The use of low-value care in Medicare is substantial. *The MedPAC Blog*, May 21. http://www.medpac.gov/-blog-/medpacblog/2015/05/21/use-of-low-value-care-in-medicare-is-substantial.

Medicare Payment Advisory Commission. 2014. *Report to the Congress: Medicare and the health care delivery system*. Washington, DC: MedPAC.

Medicare Payment Advisory Commission. 2013. *Report to the Congress: Medicare payment policy*. Washington, DC: MedPAC.

Medicare Payment Advisory Commission. 2011. *Report to the Congress: Medicare and the health care delivery system.* Washington, DC: MedPAC.

Medicare Payment Advisory Commission. 2008. *Report to the Congress: Reforming the delivery system.* Washington, DC: MedPAC.

Office of Inspector General, Department of Health and Human Services. 2016. *Nationwide analysis of common characteristics in OIG home health fraud cases.* Washington, DC: OIG.

Office of the Assistant Secretary for Planning and Evaluation, Department of Health and Human Services. 2016. *Report to Congress: Social risk factors and performance under Medicare's value-based purchasing programs.* Washington, DC: ASPE.

Rajaram, R., C. Barnard, and K. Y. Bilimoria. 2015. Concerns about using the Patient Safety Indicator–90 composite in pay-for-performance programs. *Journal of the American Medical Association* 313, no. 9 (March 3): 897–898.

Safran, D. G., M. Anastario, H. P. Rodriguez, et al. 2007. Evaluating the potential for an empirically-derived standard of performance excellence in ambulatory patient experience measures. Pre-publication draft.

Safran, D. G., D. A. Taira, W. H. Rogers, et al. 1998. Linking primary care performance to outcomes of care. *Journal of Family Practice* 47, no. 3 (September): 213–220.

Samson, L. W., L. M. Chen, A. M. Epstein, et al. 2018. Dually enrolled beneficiaries have higher episode costs on the Medicare Spending Per Beneficiary measure. *Health Affairs* 37, no. 1 (January): 86–94.

Teno, J. M., P. L. Gozalo, J. P. Bynum, et al. 2013. Change in end-of-life care for Medicare beneficiaries: site of death, place of care, and health care transitions in 2000, 2005, and 2009. *Journal of the American Medical Association* 309, no. 5 (February 6): 470–477.

Wasfy, J. H., C. M. Zigler, C. Choirat, et al. 2017. Readmission rates after passage of the Hospital Readmissions Reduction Program: A pre-post analysis. *Annals of Internal Medicine* 166, no. 5 (March 7): 324–331.

Wholey, D. R., J. B. Christianson, M. Finch, et al. 2003. Evaluating health plan quality 1: A conceptual model. *American Journal of Managed Care* 9 spec. no. 2 (June): SP53–64.