CHAPTER 5

Redesigning the Merit-based Incentive Payment System and strengthening advanced alternative payment models

CHAPTER 5

Redesigning the Merit-based Incentive Payment System and strengthening advanced alternative payment models

Chapter summary

The Medicare Access and CHIP Reauthorization Act of 2015 (MACRA) repealed the sustainable growth rate (SGR) system and established a new approach to updating payments to clinicians. It established two paths—one for clinicians who participate in advanced alternative payment models (A–APMs) and another for other clinicians (the Merit-based Incentive Payment System (MIPS)). Beginning in 2019 and continuing through 2024, clinicians on the A–APM path—that is, those who have sufficient participation in an A–APM—will receive a 5 percent incentive payment. From 2026 on, these clinicians, if they still meet the criteria for participation in an A–APM, will receive a higher update than other clinicians.

Clinicians who do not qualify for the A–APM incentive payment follow the MIPS path, which involves a separate incentive program based on clinicians' performance on certain measures. MIPS is organized into four categories (quality, cost, practice improvement, and electronic health record use), and performance in these categories determines whether clinicians receive a bonus or a penalty on their Medicare fee-for-service payments. Although budget neutral in aggregate, MIPS bonuses and penalties may have a large effect on payments for individual clinicians and hence on the attractiveness of being in an A–APM relative to MIPS.

In this chapter

As CMS has begun to implement these two paths, it is becoming apparent that there are some serious challenges. Clinicians are reporting data now for the first year of implementation for MIPS in 2019. Over 40 percent of clinicians are exempt from the program, and CMS created a very minimal standard that can be met by reporting information on one quality measure. Some stakeholders may view this approach as positive because the reporting requirements are minimal, and there will be very little effect on payment. Other stakeholders, who have invested in reporting infrastructure, may view this approach as negative. In the following years, if CMS proceeds to standards that are more difficult to meet, reporting will become more burdensome. It is not clear that the resulting data collected by CMS will be useful in detecting high and low performance, and minor differences in clinician scores could result in major differences in payment.

The implementation problems follow from basic issues in MACRA. Although MACRA repealed the SGR and attempted to address some of its shortcomings, it set up a complex system in which some signals to improve value may not be well aligned. It is always difficult mid-implementation to judge what sort of program will eventually result, but the Commission is concerned by the direction the program is taking. Therefore, although we have not made any recommendations, we have started to discuss ideas for improvement and present some of them in this chapter.

MIPS as presently designed is unlikely to succeed in helping beneficiaries choose clinicians, helping clinicians change practice patterns to improve value, or helping the Medicare program reward clinicians based on value. In part, this result is likely because the MIPS quality category is designed to allow clinicians to choose six measures from a large set of process measures, and if they choose measures that are "topped out" (measures on which everyone performs well), they will have high absolute scores. Other MIPS categories rely on clinician attestations that they are engaged in certain activities; clinicians will likely also score high on those measures. As a result, it will be difficult to ascertain any distinction among clinicians on their performance. This outcome will not be helpful to achieve the aims of MIPS, and it will impose a considerable reporting burden on clinicians. Fundamentally, it may be that individual clinicians cannot be judged on quality because there are too few cases per clinician for measures to be reliable.

This chapter discusses a possible alternative for MIPS. It starts with a quality withhold (i.e., payment rates are reduced by a set percentage and then returned or not under certain conditions) for all services paid under the physician fee schedule (PFS). It eliminates the current set of measures and instead relies on population outcome measures, such as:

- potentially preventable admissions and emergency department visits
- mortality and readmission rates
- patient experience
- healthy days at home
- rates of low-value care
- relative resource use

These measures would be calculated from claims or surveys and thus not require burdensome clinician reporting. Because these are population outcome measures, clinicians would need to be associated with populations and those populations would have to be of sufficient size for measures to be reliable. Under this construct, clinicians would need to be associated with a group of clinicians and there would be no individual-level assessment of clinician performance, only group-level assessment. Clinicians could choose to join an A–APM, join a group of clinicians that they define, elect to be measured in a group that CMS defines, or elect not to be measured at all. If not measured, they would lose the MIPS quality withhold. If in an A–APM, the withhold would be returned to them. If in a self-defined or a Medicare-defined group, their performance would be assessed as a part of the group's performance, which would determine how much of the withhold was returned or whether a quality bonus in excess of the withhold was given.

Another important aspect of MACRA is the imbalance in payment incentives for clinicians to join A–APMs or remain in MIPS. MACRA appears to encourage clinicians to join A–APMs, hence the 5 percent incentive payment for clinicians who have sufficient participation in A–APM entities. However, the design of this incentive is concerning because of potential payment inequities that could result. Under MACRA, a clinician must reach a threshold of revenue coming through an A–APM (e.g., 25 percent, 50 percent) to be eligible for the 5 percent incentive payment, and this payment is based on all of the clinician's PFS revenue, even that which does not come through an A–APM. Therefore, if the threshold for revenue coming through the A–APM is 25 percent, a practice with 24.9 percent of revenue generated through the A–APM would not be eligible for the 5 percent incentive payment, while a similar practice with 25.0 percent of its revenue through the A–APM would get a 5 percent incentive payment on all of its PFS revenue. This kind of payment cliff can introduce payment discontinuities, increase uncertainty, and appear inequitable. Therefore, we discuss making the payment reward proportional to the A–APM-generated revenue. That is, there would be no threshold and the reward would be proportional: Any revenue coming through an A–APM would secure the 5 percent payment incentive, but any other PFS revenue would not. This revision would eliminate the payment cliff and increase certainty for clinicians that their work through an A–APM entity would be rewarded.

Another aspect of balance between MIPS and A–APMs is the exceptional performance bonus available in MIPS. The bonus comes from a fund of $500 million per year (from 2019 to 2024) for clinicians with "exceptional performance" in MIPS. Moving this fund from MIPS to A–APMs would shift the incentives toward A–APMs and make MIPS less attractive. We discuss using the bonus to fund an asymmetric risk corridor for two-sided-risk accountable care organizations (ACOs) that qualify as A–APM entities.

Two-sided-risk ACOs and models like them are the A–APMs most in keeping with the Commission's principles for A–APMs discussed in the Commission's June 2016 report to the Congress. Those principles encourage A–APMs with a broader scope than some currently contemplated because the latter may lead to fragmentation, overlaps, and cross-incentives. We also discuss a possible design, in keeping with our principles, for an A–APM that could attract practices that are reluctant to take on a large amount of risk relative to their revenue.

These alternative constructs are a departure from the current design of MIPS and the application of the 5 percent A–APM payment incentive. However, they could (1) relieve clinicians of the MIPS quality reporting burden and make MIPS useful for beneficiaries, clinicians, and Medicare and (2) shift payment incentives toward greater clinician participation in A–APMs. Creating a better design for MIPS and A–APMs could help achieve Medicare's goals of improving quality for beneficiaries, making payments fair for clinicians, and restraining program costs for taxpayers. ■

Background

From 1999 to 2015, payment updates for clinicians who billed under Medicare's physician fee schedule (PFS) were covered by the sustainable growth rate (SGR) system, which set updates so that total spending would not increase faster than a target—a function of input costs, fee-for-service (FFS) enrollment, gross domestic product (GDP), and changes in law and regulation. Because annual spending generally exceeded these SGR parameters, payments to clinicians were scheduled to be reduced by ever-growing amounts starting in 2002, but the Congress overrode these negative cuts in all but the first year they were scheduled. Because of these overrides and because of volume growing in excess of per capita GDP, the resulting update reduction grew to a scheduled 21 percent in 2015.

The Medicare Access and CHIP Reauthorization Act of 2015 (MACRA) repealed the SGR system and established a new process for updating payments to clinicians. It established two paths—one for qualifying participants in advanced alternative payment models (A–APMs) and the second for all other clinicians.[1] MACRA laid out statutory updates for providers on each path.

For 2016, 2017, and 2018, updates for all clinicians under the fee schedule are 0.5 percent each year.[2] Beginning in 2019 through 2024, clinicians who meet the criteria set out in the law as qualifying APM participants receive incentive payments of 5 percent of their entire Medicare fee schedule revenue each year that they qualify.[3] From 2026 on, qualifying APM participants also receive a higher update than other clinicians: 0.75 percent versus 0.25 percent.

Under MACRA, clinicians who do not meet the A–APM criteria receive no update from 2019 through 2024 and receive lower updates than clinicians who meet the A–APM criteria in 2026 and beyond (0.25 percent). These clinicians also receive annual payment increases or decreases based on their performance in the Merit-based Incentive Payment System (MIPS), starting in 2019. Those increases and decreases in theory could be quite significant; the maximum downward adjustment increases to 9 percent of payments in 2022 and individual positive payment changes could be even greater.

The Commission commented on the proposed rule for MACRA implementation based on the discussion in its June 2016 report to the Congress (Medicare Payment Advisory Commission 2016a, Medicare Payment Advisory Commission 2016b). We noted some serious shortcomings in the MIPS program and some principles that should underlie the development of A–APMs (Medicare Payment Advisory Commission 2016b).

The final rule to implement MACRA was published on November 4, 2016 (Centers for Medicare & Medicaid Services 2016b). The final rule did not incorporate the Commission's suggestions for making MIPS a more meaningful program by focusing more on outcomes rather than process measures, and it did not follow the Commission's principles for A–APMs. Therefore, in this chapter, we present policy options for improving the design of MIPS and strengthening A–APMs. These options include redesigning MIPS to relieve reporting burden and to focus measures on outcomes of interest to beneficiaries and the program. We also address rectifying the imbalance between MIPS and A–APMs by offering a model to attract clinicians to A–APMs who are deterred from taking the risk implied in current two-sided risk models by shifting the $500 million a year (2019 to 2024) fund for clinicians with "exceptional performance" from MIPS to A–APMs, using this fund to pay for an asymmetric risk corridor for two-sided accountable care organizations that are A–APMs.

Redesigning the Merit-based Incentive Payment System

MIPS consolidates three of the existing payment adjustment programs for clinicians: the Physician Quality Reporting System (PQRS), the payment adjustment for the meaningful use of electronic health records (EHRs), and the value-based payment modifier, which includes a resource use component. The legislation allows CMS to retain the measurement process for the PQRS, EHR meaningful use, and the value-based payment modifier for use in MIPS, but merges the individual adjustments into one MIPS adjustment. MACRA continues these separate programs through 2018 and then repeals the individual programs and establishes MIPS to take effect in 2019. Under CMS's recent regulations implementing the first year of the program, clinicians must report on their quality, advancing care information, and clinical practice improvement activities during calendar year 2017 to result in a payment adjustment under MIPS that will apply in 2019.

MIPS applies to clinicians who do not qualify as A–APM participants. Annual payment increases and decreases apply based on the clinician's performance in four categories: quality, cost, clinical practice improvement

activities (CPIAs) (such as expanded practice hours), and advancing care information (ACI; formerly meaningful use of EHR). CMS has released final rulemaking for the first year of MIPS (2017 reporting year for payment adjustments in 2019) (Centers for Medicare & Medicaid Services 2016b). The first-year policies will be different from policies in later years.

MIPS assesses the first category, quality, based entirely on measures that clinicians choose to report from the MIPS measure set (based on the PQRS measure set). The roughly 275 quality measures in the MIPS measure set are largely process measures, such as whether the clinician ordered appropriate tests or followed general clinical guidelines. CMS has categorized about 170 of these measures as "high priority" because they measure outcomes (including intermediate outcome measures), patient experience, efficiency, or patient safety.

Clinicians self-attest to their performance in two other MIPS categories: CPIA and ACI. For the fourth MIPS category, cost, clinicians are assessed based on resource use (calculated from claims) relative to their peers.

Each clinician is eligible to receive a MIPS payment adjustment factor based on his or her composite performance in all four categories combined. Each clinician's composite MIPS performance score will be calculated according to weights set in law and compared against a predetermined MIPS benchmark. Clinicians above this level will receive a payment increase; clinicians below this level will receive a payment decrease.

The basic MIPS adjustments are budget neutral. MACRA set a maximum reduction for clinicians in the bottom tier of performance: 4 percent in 2019, 5 percent in 2020, 7 percent in 2021, and 9 percent in 2022 and subsequent years. The corresponding positive adjustment factors are scaled up or down to achieve budget neutrality for the basic MIPS adjustment, so the positive adjustment factors could be larger or smaller than these statutory reductions.

MACRA also appropriated an additional $500 million a year for exceptional performance in MIPS from 2019 through 2024. *Exceptional performance* is defined in the statute as performance at or above the 25th percentile above the mean (or median) of performance scores.[4]

Implementing MIPS

CMS took a "pay-for-reporting" approach for the first year of MIPS. In this approach, CMS set the MIPS benchmark at 3 points (out of 100) on the composite MIPS score, a very minimal standard. In other words, a clinician needs to score only at or above 3 points to establish eligibility for a bonus payment under MIPS in the first year. Clinicians can meet the 3-point requirement by submitting information on one quality measure, attesting to one clinical practice improvement activity, or attesting to the base advancing care information category. (CMS gave zero weight to the cost category for 2019, using its regulatory authority to override the statutory weight of 10 percent in 2019). Because of the minimal reporting requirement in the first year, CMS assumes that most MIPS-eligible clinicians (more than 90 percent) will be at or above the MIPS benchmark of 3 points. As a result, the positive payment adjustments under MIPS will be very small in the first year (Centers for Medicare & Medicaid Services 2016b).

CMS's approach for the first year of MIPS has set the administrative process in motion. As described above, in 2017, clinicians can report very little data to CMS. However, in subsequent years, clinicians may have a heavy reporting burden, and CMS will have a large amount of information to process. This information will not help CMS identify high- and low-performers, yet it could result in large differences in payment, as we discuss below.

Clinicians will be reassessed on noncomparable measures

There is wide variability in the MIPS quality measures in terms of how easy it is to achieve high performance, their relevance to the Medicare population, and their clinical relevance. Because each clinician can choose which measures to report, the amount of meaningful information received by the Medicare program varies. Under MIPS, each clinician selects six applicable measures (including an outcome measure) to report; performance on these measures determines the clinician's quality score (which is 60 percent of the MIPS score in the first year).[5] A clinician's relative performance on each measure is compared with the performance of others who reported the same measure. Many of these measures are poorly linked to outcomes of importance for beneficiaries and the program and, instead, reinforce the incentive in FFS Medicare to provide more services than are clinically necessary.

Many MIPS measures have very compressed distributions of performance. Because the measures can be reported in different ways, the result is over 600 reporting measures and method combinations for the 275 MIPS measures.[6] Of the 600, 178 are topped out (meeting CMS's criteria),

TABLE 5–1 **Potential maximum MIPS adjustments**

	2019	2020	2021	2022 and later
Base MIPS adjustments	4%	5%	7%	9%
With maximum scaling factor applied	12*	15	21	27
Plus maximum exceptional performance bonus	22*	25	31	37

Note: MIPS (Merit-based Incentive Payment System).
*Unlikely to be reached in 2019 because CMS estimates that nearly all clinicians will meet the MIPS performance standard, hence there will be very few negative adjustments to fund the positive adjustments.

Source: Medicare Access and CHIP Reauthorization Act of 2015.

and 88 have such topped-out performance that the median performance score is 100 percent.[7] For 287 measures, CMS has no performance benchmark for the first year (Centers for Medicare & Medicaid Services 2016a).

The structure of MIPS creates an inequitable system. The first inequity results from the use of self-reported quality measures, in which clinician performance is measured (and pay is adjusted) using different metrics for each clinician. The second inequity occurs because clinicians who select measures for which there is room for improvement (and that assess real, meaningful gaps in care) are much less likely to do well than clinicians who select measures on which they score highly.

Individual clinicians typically have a small number of patients qualifying for each measure

Reliably measuring performance is also a concern. For many clinicians, any individual quality measure will apply only for a subset of their patients. That number may be too small to distinguish real differences in performance on those measures from what statisticians call "noise" (unexplained variation or randomness in a sample). Combining performance on multiple measures, each with few cases, will not solve this problem.

Small differences in clinician performance may result in large differences in payment

If CMS receives compressed performance scores for quality, and two of the other three MIPS categories are attestation only, we expect that most clinicians who report to MIPS will score highly. (Those who do not report will receive the maximum negative adjustment.) In future years (when the MIPS benchmark is set at the median or mean of performance, rather than 3 points), small variations in quality measures can have an outsize effect on the MIPS composite score, even if the differences in quality performance among clinicians are clinically insignificant. Hence, payment differences may be wide (particularly if the exceptional performance bonus continues), despite the similarity of clinicians' actual performance.

The mathematical possibility for large payment adjustments in MIPS may keep some clinicians in MIPS instead of A–APMs

There is the possibility (although the likelihood is extremely small) that some clinicians could eventually receive very high payment adjustments under MIPS—up to 37 percent by 2022 (Table 5-1). This possibility arises from two factors. The first is a scaling factor to make the MIPS adjustments budget neutral: For example, if there are many more clinicians receiving penalties than bonuses, the size of the bonus would necessarily be high to maintain budget neutrality.[8] The second is the MIPS exceptional performance bonus. By statute, the MIPS exceptional performance bonus can add up to 10 percentage points to a clinician's payment adjustment.

The potential for these very high adjustments (despite the very low likelihood that they will come to pass) may provide motivation for some clinicians to remain in MIPS when they would otherwise consider joining an A–APM. CMS's MIPS APM policy, which gives participants in certain types of models high performance scores in some MIPS categories and reduces reporting burden, also works in tandem with these theoretically high MIPS payment adjustments to make MIPS relatively more attractive.

In the first year, basic MIPS adjustments will be very small for most clinicians

CMS estimates that most clinicians will receive either no adjustment or a very small positive adjustment in the first year under the basic MIPS adjustments (Centers for Medicare & Medicaid Services 2016b). CMS estimates 10 percent of clinicians will not report and will get the maximum 4 percent reduction. To preserve budget neutrality, the sum of those reductions will fund the bonuses for the other 90 percent of clinicians. Hence, the payment adjustments for the first year will be very small; CMS estimates that the maximum will be just below 1 percent (without the exceptional performance bonus). The MIPS exceptional performance bonus could add between 0 percent and 4 percent to a clinician's payment adjustment.[9]

Priorities in redesigning MIPS

MIPS, as designed, is unlikely to clearly identify high-value or low-value clinicians and hence may be of limited utility for beneficiaries (in selecting high-value clinicians), for clinicians themselves (in understanding their performance and what to do to improve), or for the Medicare program (in adjusting payments based on value).

Redesigning MIPS requires considering the current state of performance measurement and realistically setting goals for a national value-based purchasing program for clinicians. The current MIPS system is designed primarily to measure basic standards of care and processes—not outcomes. In addition, it imposes burdens on clinicians and CMS that outweigh any potential benefit because the measures used for assessing quality, the ACI category, clinical practice improvement activities, and costs are unlikely to capture true value.

Our overarching principles with respect to reforming MIPS are to measure and reward performance that is linked to outcomes and to design MIPS and A–APMs in a way that attracts a greater share of clinicians to A–APMs over time, eliminates manual clinician reporting, and develops a program that reflects the current state of performance measurement. As that state changes—for example, as data from EHRs and registries become readily available to CMS—the system should evolve to take advantage of these data.

Commission discussion: A potential redesign of MIPS

A MIPS redesign could work as follows. First, a withhold from FFS payments for all clinicians could fund a quality pool (e.g., Medicare reduces payment rates by some percentage that is sufficiently large to incentivize quality improvement). Clinicians could then:

- do nothing (and lose the withhold),
- join (or form) an A–APM (and receive the withhold back),
- join a sufficiently large group of clinicians for measurement purposes (and potentially receive a quality payment in addition to receiving the withhold back), or
- elect to be measured as part of a CMS-defined group covering a sufficiently large local population (and potentially receive a quality payment in addition to receiving the withhold back).

Under this framework, clinicians could not be worse off by choosing to be measured as a group or local area member than if they made no election at all (that is, they could not lose more than their withhold). It would also be desirable to set a maximum MIPS adjustment so that clinicians could not do better in MIPS than they could if they joined an A–APM. This redesign also contemplates moving to population-based measures rather than individual clinician-level measures. Clinicians would have the following options:

Option 1: Clinicians can choose to make no election. They would lose the withhold and would not be eligible for a quality payment. Clinicians could remain in traditional FFS and forgo any opportunity to receive a quality payment if they did not join an A–APM, join a virtual group, or elect to be measured at a local area. In other words, they would receive a reduced Medicare rate for all services (reduced by the amount of the withhold).

Option 2: Clinicians can choose to join (or form) an A–APM. Clinicians would receive their quality withhold back if they joined (or formed) an A–APM at any participation level. This option provides a modest incentive to join any A–APM and would make sure that clinicians face only one set of incentives.

Option 3: Clinicians can choose to join a "virtual" group. The virtual group, a concept introduced in MACRA (but not yet implemented through rulemaking), could mean a group of clinicians with a tax ID or legal structure in common, but could also mean a group of otherwise unrelated clinicians. For example, a virtual group could be more formally structured, such as a group practice or a group of physicians employed by a hospital,

or less formally structured, such as a physician specialty society or a geographically dispersed group of clinicians with an interest in joining together.[10]

CMS would likely have to exert some control over the size and structure of these groups to make sure the group could be measured reliably. Reliability is an issue because some clinicians are much less likely to have a sufficiently sized population of beneficiaries attributed to them. For example, a group of pathologists would be unlikely to have claims-calculated clinical outcome measures or patient experience measures, but may have relative resource use measures. CMS could set measure-specific case sizes and, in this way, implicitly require clinician groups to join with other specialties so that they would have a sufficiently large number of attributed patients for each measure.

Option 4: Clinicians can elect to be measured as part of a local or market area. CMS could define local or market areas using various characteristics. One example is to create populations of patients that use a large provider in common—for example, the hospital service area concept that groups providers together based on the hospital where their patients go most often. Under the local or market area approach, it might be possible to set a uniform case size (e.g., the local area must have at least a minimum number of beneficiaries attributed to it) so that quality measures can be robustly measured and compared against other areas or groups.

Assessing clinicians in virtual groups and local or market areas according to population-based measures (at the aggregate level)

Under a revised MIPS, CMS would use a set of CMS-calculated measures (from claims and patient experience surveys) that give insight into both the ambulatory care environment and the broader health care delivery system. Clinicians would not have to report quality data to CMS, relieving them of that burden. The Medicare program would focus on aggregate measures extracted from claims that assess care for patients across the continuum of providers, such as:

- potentially preventable admissions and emergency department visits
- mortality and readmission rates after inpatient hospital stays
- healthy days at home
- patient experience
- rates of low-value care
- relative resource use

These measures are intended to be illustrative; in general, the goal would be to use claims- and survey-calculated measures that assess performance in the categories of clinical outcomes, patient experience, and efficiency. In this redesign, MIPS would no longer include clinical practice improvement activities and EHR technology as separate categories requiring clinician attestation.[11] In addition, even clinicians who elected group- or area-level measurement would not be required to report any quality measures to CMS.

Changing the focus to assessing population-based outcomes

The alternative design described above incorporates some trade-offs, by necessity. The key one is that the Medicare program would no longer score an individual clinician's performance and no longer require clinician reporting. The concept is to adopt a broader, claims- and survey-calculated uniform measure set that assesses the overall performance of a health care delivery system and its clinicians. These population-based measures are generally not reliable at the individual clinician level. The Medicare program would assess performance (and adjust payment) based only on performance at a group or local area level. Clinicians could elect not to receive a quality payment, but if they wished to be eligible for a quality payment, they would need to join (either actively or passively) a set of clinicians to be measured (or move to an A–APM and be eligible to get back their quality withhold).

The benefits of using population-based measures are significant. First, this approach sends clinicians a signal that they should view the care they provide as part of a continuum that crosses sectors and incorporates the totality of patient care. This perspective helps to counter the silo-driven FFS system that encourages providers to focus only on the services they directly provide. Second, it aligns with other programs in Medicare (such as the Commission's vision for comparing quality across Medicare Advantage, FFS, and accountable care organizations (ACOs) (Medicare Payment Advisory Commission 2014b), sending the same set of signals to all providers involved. Third, it keeps Medicare's focus on broad, aggregate measures of performance and leaves it to provider entities (hospitals, health systems, ACOs) to determine how best to measure and assess quality in their particular environment.

Fourth, it reduces practice cost and burden on clinicians by eliminating all clinician reporting of measures.

There are drawbacks to such a redesign. CMS is already years down the path of establishing a comprehensive quality-data reporting system that uses multiple methods of data reporting and extraction. CMS has modified this system to support MIPS as well as the two additional MIPS categories that clinicians must report (advancing care information and clinical practice improvement activities). Switching gears at this point would require significant time and effort for CMS. In addition, clinicians and other providers in the broader health care delivery system have spent significant time and resources building systems and operations that feed information to CMS using this framework.

Because it would measure clinician performance at a group or regional level, the potential MIPS redesign would not help beneficiaries choose a clinician who meets their preferences—for example, a surgeon with low complication rates or a primary care clinician with good improvement in patient function. A separate issue, not discussed in this chapter, is the use of quality information for public reporting. In this chapter, we are concentrating on MACRA as it affects clinician payment—which is complex enough.

Furthermore, providers may feel that population-based outcome measures do not reflect their individual performance, and because the measurement would be group based or regional, it reflects care that is outside their control. The potential redesign would require population-based outcome measures; appropriate risk adjustment; and policy decisions about the amount of the withhold, the allocation of bonus dollars among groups, and the form and amount of the quality payment.

Despite these challenges, it is worth recalling the status quo. Presently, CMS collects a large amount of information using a variety of sources, with varying clinician burden and varying value. However, nearly half of the measures have compressed performance, and many of them measure minimal standards of care. CMS does not presently use them for public reporting through Physician Compare, in part because of the inability to compare across all providers and small sample sizes. Individual-level quality measurement is inherently challenging. Measurement at the group level can be more reliable but does not provide information on individual clinicians. This tension will not be resolved under any design. CMS has delayed full implementation of MIPS for one year, but will still face these problems in the future.

In the future, as EHRs and registries mature and become more interoperable, it might be possible to overcome some of the current limitations of quality measurement for clinicians. At that point, it might be possible for the Medicare program to assess clinician performance more readily using sources other than claims and surveys (such as EHRs or clinician data registries). However, given the current state of the art of quality measurement and the lack of interoperability (and possible data blocking) between EHRs, the design for MIPS is not now tenable.

One outcome of a redesign such as the one above is that clinicians could see signals to join an organized entity that assumes responsibility for the cost and quality of patient care. For example, if clinicians would like to receive a quality payment but do not like being measured against the performance of their local area, they could seek a group (a virtual group either more or less formal) with which to be measured. This option could prepare them to transition more easily to a structure like an ACO or other A–APM. The downside is that it could create further incentives for provider consolidation, which can increase Medicare and private-sector spending (see Chapter 10 of this report).

Rectifying the imbalance between MIPS and A–APMs

If MACRA is intended to move clinicians toward participating in A–APMs (as evidenced by the 5 percent incentive payment and higher updates in later years for clinicians participating in A–APMs), certain aspects of the law and its implementation may undermine this intent. Those aspects could make remaining in MIPS too attractive relative to A–APM participation or could make the benefits of participating in A–APMs too uncertain. Below, we discuss two policies that could help rectify this imbalance. We do not endorse policies that reward simply being in an A–APM or make it easier for an A–APM to appear to succeed; those policies undermine the concept of alternative payment models that further delivery system reform. Instead, the principles we developed last year emphasize the development of A–APMs with the potential to improve care coordination for patients over the entire course of care while protecting the Medicare program and taxpayers from excessive spending (Medicare Payment Advisory Commission 2016b). The less restrictive

definition of A–APMs that some put forward might make A–APMs more available and might make it easier for them to appear to succeed but would not necessarily result in A–APMs that further the goals of the Medicare program as the Commission understands them.

Applying the A–APM incentive payment to clinicians' revenue coming through the A–APM

Under MACRA, the 5 percent A–APM incentive payment is applied to a clinician's entire Medicare physician fee schedule (PFS) revenue from the prior year. However, to qualify for the incentive payment, a clinician (or, as defined in regulation, an A–APM entity) must meet the threshold for the share of PFS revenue coming through an A–APM. That numerical threshold is set in statute and increases over time. In 2019 and 2020, a clinician practice must have at least 25 percent of its PFS revenue through an A–APM, 50 percent in 2021 and 2022, and 75 percent in 2023 and later. Uncertainty about meeting this threshold could deter clinician participation in A–APMs.

We consider an alternative policy under which there would be no numerical threshold for participation, and instead, the 5 percent A–APM incentive payment would apply only to PFS revenue coming through the A–APM rather than to all of a clinician's PFS revenue. That is, the policy would make the incentive proportional to involvement in the A–APM. This approach would greatly simplify administration of the policy, increase the certainty of a reward for moving services into A–APMs, and make the policy fairer to clinicians. For example, it would avoid the situation of a clinician practice with 24.9 percent of PFS revenue coming through an A–APM receiving no incentive payment, and one with 25.0 percent of revenue coming through the A–APM getting a 5 percent incentive payment on all of its PFS revenue.

Under this alternative, the incentive would depend solely on the revenue of the practice that comes through the A–APM, which means that any work done through an A–APM would be rewarded with certainty. In addition, there would be no payment cliffs or discontinuities at the thresholds. (Additionally, such a revised design would help avoid uncertainty for practices that may be concerned they will lose the incentive payment as the threshold rises from 25 percent, to 50 percent, to 75 percent in later years.)

The alternative would also reduce administrative complexity. Under current policy, CMS first calculates the ratio of the entity's PFS revenue through the A–APM and its total PFS revenue. If that ratio falls short of the threshold, CMS then calculates a "patient-count ratio"—the ratio of patients attributed to the A–APM and the practice's total patients—to determine whether that ratio meets the threshold. CMS has proposed different (lower) thresholds for the patient-count method (Centers for Medicare & Medicaid Services 2016b). In addition, MACRA has an "all-payer" option in later years that requires CMS to determine what share of revenue or patients is coming through A–APM-like arrangements for other payers. That determination could require access to a practice's contracts with other payers and could be a large administrative burden on all parties. The alternative policy, which eliminates the revenue threshold, would make the patient-count and all-payer calculation methods unnecessary.

Under MACRA, clinicians are exempt from MIPS if they meet the numerical threshold (e.g., 25 percent of PFS revenue comes through an A–APM). Because the alternative policy would have no numerical threshold, determining which clinicians were exempt from MIPS would require different parameters. Under the MIPS policy option described earlier, clinicians with any A–APM participation would be exempt from MIPS, and their quality withhold would be returned to them.

Revising the model to encourage taking on two-sided risk

MACRA was designed to encourage clinicians to participate in A–APMs that place them at more than nominal financial risk. In part, this design may have been chosen because incentives to achieve savings are stronger in properly structured models with two-sided risk (i.e., there is a reward for reducing spending below a benchmark and a penalty for exceeding a benchmark) than in one-sided models, which have no penalty if spending exceeds a benchmark. In addition, a two-sided risk model provides some protection for the Medicare program from losses and could allow CMS to waive certain regulations designed to protect against overuse of services (Medicare Payment Advisory Commission 2014a). At the same time, MACRA is a clinician-focused policy that addresses payments for clinicians and creates incentives for them to join certain models. Thus, when considering a redesign of MACRA, this chapter focuses on two-sided risk models that clinicians might consider attractive.[12]

In addition, the Commission maintains that a principle for A–APMs is that the entity should be at financial risk for total Part A and Part B spending (Medicare Payment

TABLE 5–2 **Illustrative comparison of benchmark-based and revenue-based risk**

1,000	Number of beneficiaries
$10,000	Per capita Part A and Part B benchmark
$10,000,000	Total Part A and Part B benchmark
$300,000	**Benchmark-based standard: 3 percent of benchmark**
$500,000	Practice revenue through the ACO (assumed to be 5 percent of Part A and Part B)
	Revenue-based standard: 8 percent of total FFS practice revenue
$40,000	Low end: Total practice revenue is $500,000, all comes through A–APM
$160,000	High end: Total practice revenue is $2,000,000, 25 percent comes through A–APM

Note: ACO (accountable care organization), FFS (fee-for-service), A–APM (advanced alternative payment model). We assume that the only ACO participants are clinicians, and they are accountable for all Part A and Part B spending for the year.

Advisory Commission 2016b). This principle is directed at two goals: (1) to achieve the clinical and financial integration promised by a reformed payment system and (2) to reduce the risk of excess spending without value. However, one issue in making two-sided risk models accessible to a clinician group is that taking risk under a Part A and Part B benchmark might make the downside risk look too formidable to attempt. For example, there is usually a large difference between a clinician group's revenue through an ACO and its ACO's total Part A and Part B spending benchmark. Although clinicians influence a large share of Medicare spending, spending under the PFS itself is about 15 percent of total Medicare spending; most spending goes to other providers. In addition, a physician group would be very unlikely to capture all PFS spending as revenue for its attributed beneficiaries. A primary care group's revenue through an ACO would likely account for only about 5 percent of the Part A and Part B benchmark. Thus, benchmark spending in an ACO would be a large multiple of a clinician group's revenue through the ACO. That multiplier would be advantageous if the practice is in a one-sided risk model, but it could seem too much to venture if the practice was at two-sided risk for total spending.

One approach to resolving this dilemma is to limit the risk for the clinicians' practice. The law requires that an A–APM be at more than nominal risk, and CMS has established two options for a nominal-risk standard: either a benchmark-based standard (3 percent of the model's benchmark) or a revenue-based standard (8 percent of an entity's FFS revenue) (Centers for Medicare & Medicaid Services 2016b). In general, the benchmark-based standard represents more risk for a clinician practice than the revenue-based standard.[13]

To illustrate these differences, we consider the case of a two-sided-risk ACO and demonstrate that the revenue-based risk standard will be less than the benchmark-based standard. In this example, assume that the only participants in the ACO entity are clinicians, that they are accountable for all Part A and Part B spending for the year, and that the ACO has 1,000 beneficiaries attributed to it.[14] Also, assume the benchmark per capita Part A and Part B spending is $10,000. CMS set a 3 percent benchmark-based standard for nominal risk or an 8 percent revenue-based standard.[15]

Under these assumptions, the spending benchmark for the entity would be $10,000,000, and 3 percent of that would be $300,000 (the benchmark-based standard) (Table 5-2).

For CMS's revenue-based standard in this example, we assume that the ACO entity (which we will refer to as the practice) has Medicare FFS revenue coming through the ACO equal to 5 percent of the benchmark, or $500,000. CMS would require a minimum risk of 8 percent of the

practice's total FFS revenue. In this example, the practice's total revenue could range from $500,000 to $2,000,000. Total practice revenue must be at least $500,000—the amount coming through the ACO. The most its total revenue could be is $2,000,000—because, at a minimum, 25 percent must come through the ACO to meet the threshold, and 25 percent of $2,000,000 is $500,000.

Hence, 8 percent of total revenue must range between $40,000 (8 percent of $500,000) and $160,000 (8 percent of $2,000,000). In this example, both the minimum ($40,000) and the maximum ($160,000) amounts at risk in the revenue-based standard are less than the $300,000 at risk under the benchmark-based standard. Therefore, CMS's 8 percent of practice-revenue standard would represent less risk for the practice than the 3 percent of benchmark standard.[16]

Next, we describe a revised model in which revenue is defined as the practice's Medicare revenue coming through the A–APM (instead of CMS's definition of all Medicare practice revenue). Under this example, the 8 percent limit of the amount at risk would be $40,000 (8 percent of $500,000). This revised policy could encourage clinician groups to participate in A–APMs with more than nominal risk because it would represent a lower level of risk for the practice than the benchmark-based standard ($300,000 for the illustrative ACO model in Table 5-2) and would be the low end of CMS's revenue-based standard. This definition would be consistent with the revised 5 percent incentive payment discussed earlier. That is, the 5 percent incentive payment is proportional, applying only to the practice's revenue coming through an A–APM.

The effective risk for the practice would thus be even lower because of the 5 percent incentive payment. After accounting for the 5 percent incentive payment, the effective risk would be 3 percent of the practice's revenue coming through the A–APM (8 percent minus 5 percent). In the example in Table 5-2, that effective risk would be $15,000 (3 percent of $500,000).[17]

Thus, a revised model could:

- define revenue in the revenue-based standard as a practice's Medicare FFS revenue coming through the A–APM—consistent with the proposal to compute the 5 percent incentive on revenue through the A–APM.
- have a revenue-based instead of a benchmark-based nominal risk standard. (For consistency, the model could also define the top as well as the bottom of a risk corridor—the limit for savings and losses—in Medicare revenue terms.) (See the following section for further discussion of risk corridors.)

Consistent with the Commission's principles, shared savings and losses would be based on total Part A and Part B performance (while limited by a risk corridor), and small entities would need to aggregate to reliably detect cost and quality performance.[18] The intent is to create an incentive that is large enough to motivate improvement but limit the loss to something a practice could reasonably take on.

Retargeting the MIPS "exceptional performance" fund

MACRA appropriated an additional $500 million a year for "exceptional" performance in MIPS. This payment goes to any clinician at or above the 25th percentile above the MIPS performance standard, and the exceptional performance bonus is proportional. We have pointed out that the distribution of scores in MIPS may be very tight, with little real distinction between relatively high and low scores because almost all clinicians who report could have a very high absolute score. As a result, the MIPS exceptional performance bonus payments could be distributed to clinicians whose performance is essentially equivalent to those who do not get the bonus (e.g., those who score 99.8 versus those who score 99.6). In addition, in later years, the budget-neutral MIPS adjustments could give substantial rewards to the top scorers. Adding to this reward could theoretically create such a large reward that it would discourage clinicians from moving from MIPS to A–APMs.

One policy option would be to eliminate the $500 million MIPS exceptional performance bonus (so that MIPS becomes budget neutral) and return it to the Treasury or retarget the money. We discuss a retargeting option below that takes the revenue from the fund and uses it to help entities in A–APMs move toward two-sided risk by funding asymmetric risk corridors in two-sided-risk ACOs.[19]

A risk corridor limits the amount of savings or losses for which an entity is at risk. For example, if an entity's revenue through an ACO were $500,000, a 20 percent risk corridor would mean that the most the entity could gain or lose in shared savings or shared losses would be $100,000 (see Column 1 of Table 5-3, p. 172).[20] An asymmetric risk corridor could decrease the amount at risk, increase the maximum amount on the upside, or do both. Table 5-3 shows an example (Column 2) that increases the upside—

TABLE 5–3 **Illustration of symmetric and asymmetric risk corridors in two-sided-risk ACOs**

	Risk corridor for a clinical group with $500,000 of revenue through the ACO	
	Symmetric +20 percent / –20 percent of revenue	**Asymmetric +100 percent / –20 percent of revenue**
Limit on shared savings	$100,000	$500,000
Limit on shared losses	–$100,000	–$100,000

Note: ACO (accountable care organization).

the amount of shared savings allowed. Building on the example above, we compare illustrative symmetric and asymmetric risk corridors.

In the illustrative example in Table 5-3, the upper and lower risk corridor in the symmetric case are set at +/–$100,000, 20 percent of the clinician group's $500,000 in Medicare revenue through the ACO. In the asymmetric example, the upper limit on the risk corridor is 100 percent of revenue, or $500,000, and the lower limit is $100,000. The percentages in Table 5-3 are purely illustrative.

Effectively, the maximum reward would also include the 5 percent A–APM incentive, which would be $25,000 in this example (5 percent of the $500,000 in revenue through the A–APM). Thus, the upper limit on the practice's reward in the asymmetric case would be effectively $525,000. If the 5 percent incentive were paid on all revenue through the A–APM regardless of success in the A–APM, the loss would be at most $75,000 in both the symmetric and asymmetric cases in this example.

Additional money would be needed to fund asymmetric risk corridors because some ACOs would get shared savings and some would get shared losses from random variation. If the risk corridor were symmetric, savings and losses from random variation would balance out over the years from the Medicare program's perspective. However, if the risk corridor had higher upper than lower limits, Medicare could expect to pay out more in unwarranted shared savings than it would collect in unwarranted shared losses, overall. Because the additional spending is a potential liability for the program, one option is to retarget the $500 million in funds designated to reward exceptional performance under MIPS. The total funding needed would have to be estimated, which would require knowing the number of two-sided-risk ACOs eligible, the number of beneficiaries in each, their benchmarks, and the revenue of the clinicians coming through the ACOs. Random variation decreases as the attributed population increases, and that would also need to be factored into the calculation. The asymmetric risk corridor model would be transitional because it would terminate at the end of 2024 along with the funding for the MIPS exceptional performance bonus.

The model is designed to selectively attract clinician groups because the revenue-based standards are designed for groups whose revenue through the ACO is a small share of the total benchmark (Part A and Part B) spending. Performance would continue to be judged against total Part A and Part B spending. Hospital-based ACOs would tend toward models with a benchmark-based standard with higher benchmark-based rewards because their share of the benchmark spending would tend to be higher than a clinician group's share. Essentially, as an entity's revenue as a share of the benchmark increases, revenue-based and benchmark-based standards would converge. As an ancillary benefit, this model would likely indirectly provide support to primary care providers (PCPs). It would reward PCPs to the extent that attribution to the ACO is based on primary care evaluation and management claims, the extent that better primary care leads to savings in Medicare spending, and the extent that ACOs pass on rewards to primary care clinicians.

Conclusion

MACRA and its implementation has created a complex system that will not identify or appropriately reward high- and low-value clinicians, requires a massive reporting effort, and sends conflicting signals as to which models clinicians should move to. The Commission is concerned by the direction the program is taking in its first year and, although it is always difficult mid-implementation to judge what sort of program will eventually result, there appear to be basic aspects of the program that will make it difficult for it to succeed in later years. Therefore, although the Commission has not made any recommendations, we have introduced in this chapter three possible options to further policy discussions.

First, an alternative design could eliminate reporting burden and create incentives for clinicians to move to high-value models. MIPS as now designed will place a heavy burden on providers and CMS, but it is unlikely to identify high-value clinician performance. One potential redesign would reorient MIPS toward assessing the performance of groups of clinicians on population-based outcome measures.

Second, a modification of the 5 percent A–APM incentive payment could simplify the system and increase equity by applying the 5 percent A–APM incentive payment only to clinicians' revenue through the A–APM.

Third, to address the relative attractiveness of MIPS versus A–APMs, the MIPS exceptional performance bonus fund could be used to finance support for A–APMs. One way to do so would be to establish a two-sided-risk ACO model that contains an asymmetric risk corridor, allowing the upside to be greater than the downside risk. Further, the downside risk could be limited to a share of clinician revenue through the ACO. This approach would give clinician groups a path to two-sided risk that they might find attractive.

These options are meant to inform further policy discussions and to start to address the inherent difficulties in assessing clinician performance and the challenges of moving clinicians toward reformed payment and delivery systems. ■

Endnotes

1 For clarity, we use the terms CMS created and uses in the final rule: for example, *A–APM* instead of *eligible alternative payment model*, the term used in the statute.

2 Other policies in statute may affect the fee schedule payment update in any given year. For example, CMS did not achieve a required level of savings resulting from identifying misvalued codes, and so the effective update in 2016 was less than 0.5 percent.

3 The statute and regulation define the clinicians receiving the 5 percent incentive payment as "qualifying APM participants."

4 If the mean or median MIPS score is 50 points and performance scores are equally distributed, then all clinicians with a score at or above 67.5 points will receive a MIPS exceptional performance bonus, and the MIPS exceptional performance bonus will increase linearly from 67.5 points to the maximum performance score.

5 In the first year, the weighting is 60 percent quality, 15 percent CPIA, 25 percent ACI, and 0 percent cost. By 2021, the weighting is 30 percent quality, 15 percent CPIA, 25 percent ACI, and 30 percent cost. *Applicable* is defined as measures relevant to a particular MIPS-eligible clinician's services or care rendered. CMS has identified 26 specialty measure sets (e.g., cardiology, allergy/immunology, internal medicine) to help clinicians identify applicable measures. Clinicians can receive bonus points for reporting "high-priority" outcomes, patient experience, efficiency measures, or patient safety measures. Clinicians also have the option to report more than six measures and have CMS choose the six that give the best result.

6 The present methods of MIPS reporting are administrative claims, claims, Consumer Assessment of Healthcare Providers and Systems® for MIPS, CMS web interface, EHRs, registry, or Qualified Clinical Data Registry.

7 CMS calculates the performance of all other clinicians who reported the same measure using the same reporting mechanism (e.g., all clinicians that reported a bariatric screening measure using a registry). In its final rule for the 2019 payment year., CMS described various proposals for dealing with topped-out measures and may propose changes to the scoring for topped-out measures in the 2020 rule (that correspond to 2018 quality measure reporting).

8 This scaling effect could occur, for example, because CMS will set the benchmark prospectively. Actual performance may vary.

9 This estimate assumes that the number of clinicians (and their associated Medicare revenue) is evenly distributed above and below the MIPS exceptional performance threshold and that the MIPS exceptional performance threshold is set at 25 percent above the median of performance scores.

10 Some large group practices may have enough clinicians for reliably assessing population-based measures.

11 Assessing patient experience of care by surveying patients directly could give a truer picture of clinical practice improvement, such as greater continuity, after-hours access to needed services, and whether clinicians help facilitate transitions across providers and settings. Currently, the CPIA category in MIPS requires only that the clinician attest that they adopted these processes, even though the processes may not translate into meaningful changes for patients.

12 In theory, on the one hand, clinician practices may be well positioned to achieve savings under an A–APM model because in most cases they do not lose their own FFS revenue if they reduce services such as emergency department visits, inpatient admissions, and post-acute care use. Hence, their incentive to reduce such services may be greater than an A–APM with hospitals as participants. On the other hand, a system that includes hospitals as well as clinicians may control a broader span of services and be better able to coordinate care.

13 For entities that include hospitals as well as clinicians (i.e., the more services provided through the entity), the benchmark-based and revenue-based standards might start to converge because the revenue for the entities would include more of the benchmark.

14 We use 1,000 attributed beneficiaries for ease of illustration only. Medicare Shared Savings Program ACOs for example, must have over 5,000 attributed beneficiaries.

15 These are the minimum standards. Individual models can have higher standards.

16 For entities that have both clinicians and hospitals as participants, the revenue-based and benchmark-based standards would start to converge as the entity's revenue through the A–APM accounted for a larger share of the benchmark.

17 Policymakers would have to decide on the magnitude of the loss limit. Although 8 percent is the current standard for more than nominal risk, individual models have higher limits. CMS is considering raising the minimum in future

years. The 8 percent revenue standard is in effect for the 2017 and 2018 qualified practitioner (QP) performance periods. (The 2017 QP performance period will be used to determine which clinicians are QPs for 2019.) It is not defined for 2019 and after, but two possibilities are offered: 15 percent of revenue or 10 percent of revenue so long as risk is at least equal to 1.5 percent of benchmark.

18 Those principles are discussed in our June 2016 report to the Congress. They include making incentive payments only if the A–APM entity were successful in controlling cost, improving quality, or both; holding an A–APM entity at risk for total Part A and Part B spending; holding the entity responsible for a beneficiary population sufficiently large to detect changes in spending or quality; giving the entity the ability to share savings with beneficiaries; and having CMS give the entity regulatory relief. As discussed in our comment letter on MACRA implementation, some of the proposed A–APMs (e.g., two-sided-risk ESRD (end-stage renal disease) Seamless Care Organizations) are consistent with those principles and others (e.g., Comprehensive Primary Care Plus) are not (Medicare Payment Advisory Commission 2016a, Medicare Payment Advisory Commission 2016b).

19 We discuss two-sided-risk ACOs because they (and models like them) are the A–APMs that most closely align with the Commission's principles for A–APMs.

20 It should be noted that the practice would likely have Medicare FFS revenue outside the ACO that would not be at risk, thus the amount at risk would be a smaller share than 20 percent of the practice's total FFS revenue.

References

Centers for Medicare & Medicaid Services, Department of Health and Human Services. 2016a. Alternative payment models in the Quality Payment Program. https://qpp.cms.gov/docs/QPP_Advanced_APMs_in_2017.pdf.

Centers for Medicare & Medicaid Services, Department of Health and Human Services. 2016b. Medicare program; Merit-based Incentive Payment System (MIPS) and Alternative Payment Model (APM) Incentive under the physician fee schedule, and criteria for physician-focused payment models. Final rule. *Federal Register* 81, no. 214 (November 4): 77008–77831.

Medicare Payment Advisory Commission. 2016a. Comment on CMS's proposed rule on the Merit-based Incentive Payment System and alternative payment models, June 15. http://www.medpac.gov/docs/default-source/comment-letters/medpac-comment-on-cms-s-proposed-rule-on-the-merit-based-incentive-payment-system-and-alternative-pa.pdf?sfvrsn=0.

Medicare Payment Advisory Commission. 2016b. *Report to the Congress: Medicare and the health care delivery system.* Washington, DC: MedPAC.

Medicare Payment Advisory Commission. 2014a. Comment letter to CMS on accountable care organizations, June 16. http://medpac.gov/docs/default-source/comment-letters/comment-letter-to-cms-on-accountable-care-organizations-june-16-2014-.pdf?sfvrsn=0.

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