



CENTER FOR
HEALTHCARE
QUALITY &
PAYMENT REFORM

How to Create an Alternative Payment Model

*Designing Value-Based Payments That Support
Affordable, High-Quality Healthcare Services*

Harold D. Miller



First Edition
December 2018

CONTENTS

EXECUTIVE SUMMARY	1
I. WHAT IS AN ALTERNATIVE PAYMENT MODEL?	1
A. The Need for Alternative Ways of Paying for Healthcare Services	1
B. How MACRA Defines an Alternative Payment Model	1
C. Physician-Focused APMs, Advanced APMs, and MIPS APMs	4
D. The Need for More and Better APMs	5
II. OVERVIEW OF HOW AN APM IS CREATED	7
A. Identifying Opportunities for Achieving Savings and Improving Quality	7
B. Identifying Needed Changes in Care Delivery	8
C. Identifying the Barriers in the Current Payment System	8
D. Designing the APM	9
E. Operationalizing the APM	9
F. Implementing the APM	9
III. IDENTIFYING OPPORTUNITIES FOR SAVINGS AND QUALITY IMPROVEMENT	11
A. Opportunities for Reducing Spending on Planned Care	12
B. Opportunities for Reducing Spending on Unplanned Care	14
C. Opportunities for Improving Quality/Outcomes Without Savings	16
D. Avoiding Increases In Spending	17
IV. IDENTIFYING CHANGES IN CARE NEEDED TO IMPROVE CARE	19
A. Identifying How Services Will Need to Change	19
1. Time and Costs of Doing Less	19
2. Ensuring Availability of Alternative Services	20
3. Planned Services Needed to Reduce Unplanned Care	20
B. Determining the Costs of the Services to Be Delivered	22
1. Developing a Cost Model for Services	22
2. Identifying Startup Costs	25
3. Considering the Cost of Time	25
4. Determining Whether Costs Differ for Different Patients	25
5. The Challenges of Estimating Costs Before Services Are First Delivered	25
C. Defining the Business Case for an Alternative Payment Model	26
1. Ensuring There is a Business Case for Both Payers and Providers	26
2. Addressing Uncertainty in the Business Case	26
V. IDENTIFYING THE BARRIERS IN THE CURRENT PAYMENT SYSTEM	28
A. Lack of Payment for Services	28
B. Underpayment for Services	30
C. Inability to Control Services Delivered by Other Providers	32
D. Barriers Created by Cost-Sharing	32
E. Other Barriers	33

VI. DESIGNING THE APM.....	34
A. APM Component #1: Removing the Barriers in the Current Payment System	34
1. Paying for Unpaid Services	35
2. Aligning Payments With the Costs of Services.....	37
3. Enabling Control of Services Delivered by Other Providers	52
4. Modifying Cost-Sharing	54
B. APM Component #2: Creating Accountability for Spending.....	57
1. Defining the Accountability Measures.....	58
2. Setting the Performance Targets for Utilization and Spending.....	66
3. Assessing Performance on Utilization/Spending	76
4. Making Performance-Based Adjustments to Payments.....	81
C. APM Component #3: Creating Accountability for Quality	95
1. Defining the Accountability Measures.....	96
2. Setting the Performance Targets for Quality.....	101
3. Assessing Performance on Quality	110
4. Making Performance-Based Adjustments to Payments.....	111
D. APM Component #4: Defining the Eligible Patients	120
1. Establishing Appropriate Eligibility Criteria	120
2. Prospective vs. Retrospective Eligibility Determination	121
3. Avoiding Cherry-Picking and Lemon-Dropping.....	122
4. Patient Agreements as a Component of Eligibility Criteria.....	124
E. Finalizing the APM Design	125
1. Combining the Components of the APM	125
2. Ensuring the APM Design Supports the Business Case.....	138
3. Creating Alternative Versions of the Alternative Payment Model	139
VII. OPERATIONALIZING THE APM	141
A. Changing Payments for Services	141
B. Determining Eligibility of Patients	152
C. Measuring Performance	152
D. Adjusting Payments for Performance	155
1. Penalties/Bonuses Based on Achieving Targets	155
2. Outcome-Based Payments.....	157
3. Bundled/Warrantied Payments.....	157
VIII. IMPLEMENTING THE APM	158
A. Obtaining Participation by Payers, Providers, and Patients	158
1. Encouraging Participation by Payers	158
2. Encouraging Participation by Providers.....	162
3. Encouraging Participation by Patients	167
B. Finalizing the APM Parameters.....	169
1. Challenges in Finalizing the Initial Parameters for the APM	169
2. Using a Beta Testing Process for APMS	170
C. Evaluating the APM	171
1. Defining What Should Be Evaluated	171
2. Defining the Timeframe for Evaluation	172
3. Setting Thresholds for Success	172
4. Taking Actions Based on the Evaluation Results	173
5. Using Better, Faster Approaches to Evaluating APMS.....	176
D. Revising the APM Parameters	177
1. The Need to Adjust APM Parameters Over Time	177
2. Alternative Approaches to Updating Parameters	178
3. Issues in Updating APM Parameters	181

IX. CREATING BETTER VALUE-BASED PAYMENT MODELS	184
A. Correcting the Problems with Fee for Service Payment.....	184
B. Preserving the Strengths of Fee-for-Service Payment.....	185
C. The Need for Multiple Alternative Payment Models	186
D. The Need for Payment Reforms Other Than APMs.....	186
E. Accelerating Progress Toward More Affordable Care.....	186
APPENDIX: EXAMPLES OF ALTERNATIVE PAYMENT MODELS.....	187
APM #1: PAYMENT FOR A HIGH-VALUE SERVICE.....	188
1. Opportunity for Savings and Quality Improvement.....	188
2. Changes in Care Delivery Needed and Associated Costs	188
3. Barriers in the Current Payment System.....	189
4. Design of the APM	191
5. Operationalizing the APM.....	194
6. Implementing the APM.....	195
APM #2: TREATMENT OF AN ACUTE CONDITION.....	197
1. Opportunities for Savings and Quality Improvement.....	197
2. Changes in Care Delivery Needed and Associated Costs	197
3. Barriers in the Current Payment System.....	199
4. Design of the APM	200
5. Operationalizing the APM.....	206
6. Implementing the APM.....	207
APM #3: MANAGEMENT OF A CHRONIC CONDITION	209
1. Opportunities for Savings and Quality Improvement.....	210
2. Changes in Care Delivery Needed and Associated Costs	211
3. Barriers in the Current Payment System.....	211
4. Design of the APM	212
5. Operationalizing the APM.....	220
6. Implementing the APM.....	222
ENDNOTES	224

EXECUTIVE SUMMARY

WHAT IS AN ALTERNATIVE PAYMENT MODEL?

There is broad consensus that fee-for-service payment is a major reason why healthcare spending has grown faster than inflation without any corresponding improvement in the quality of care for patients. To address this, the Medicare Access and CHIP Reauthorization Act of 2015 (MACRA) authorized the creation of “Alternative Payment Models” in Medicare. In general, an APM must either:

- improve the quality of care without increasing spending;
- reduce spending without reducing the quality of care; or
- improve the quality of care and reduce spending.

As of 2018, the majority of healthcare providers in the country were not participating in an Alternative Payment Model, and most providers had not even had an opportunity to do so because of the small number and narrow focus of the APMs that had been created. Moreover, the APMs that do exist have generally failed to achieve any significant savings.

Although many people believe the poor performance of current APMs is because they do not create enough “financial risk” for the participating providers, there is no evidence that simply increasing financial risk would result in greater savings. On the other hand, transferring financial risk to providers can have undesirable results, including loss of access to services for higher-need patients, higher prices due to consolidation of providers, and lower quality of care.

A more plausible explanation for the failure of current APMs is that the APMs have not actually solved the problems with fee-for-service payment. For example, most APMs do not actually change the underlying fee-for-service system, but simply provide bonuses to healthcare providers when spending is reduced.

Fortunately, there are different and better ways to design Alternative Payment Models that can directly address the problems in the fee-for-service system without placing healthcare providers at significant financial risk or causing patients to worry about whether needed care is being withheld for financial reasons.

HOW TO CREATE A SUCCESSFUL ALTERNATIVE PAYMENT MODEL

Creating a successful APM requires a six-step process:

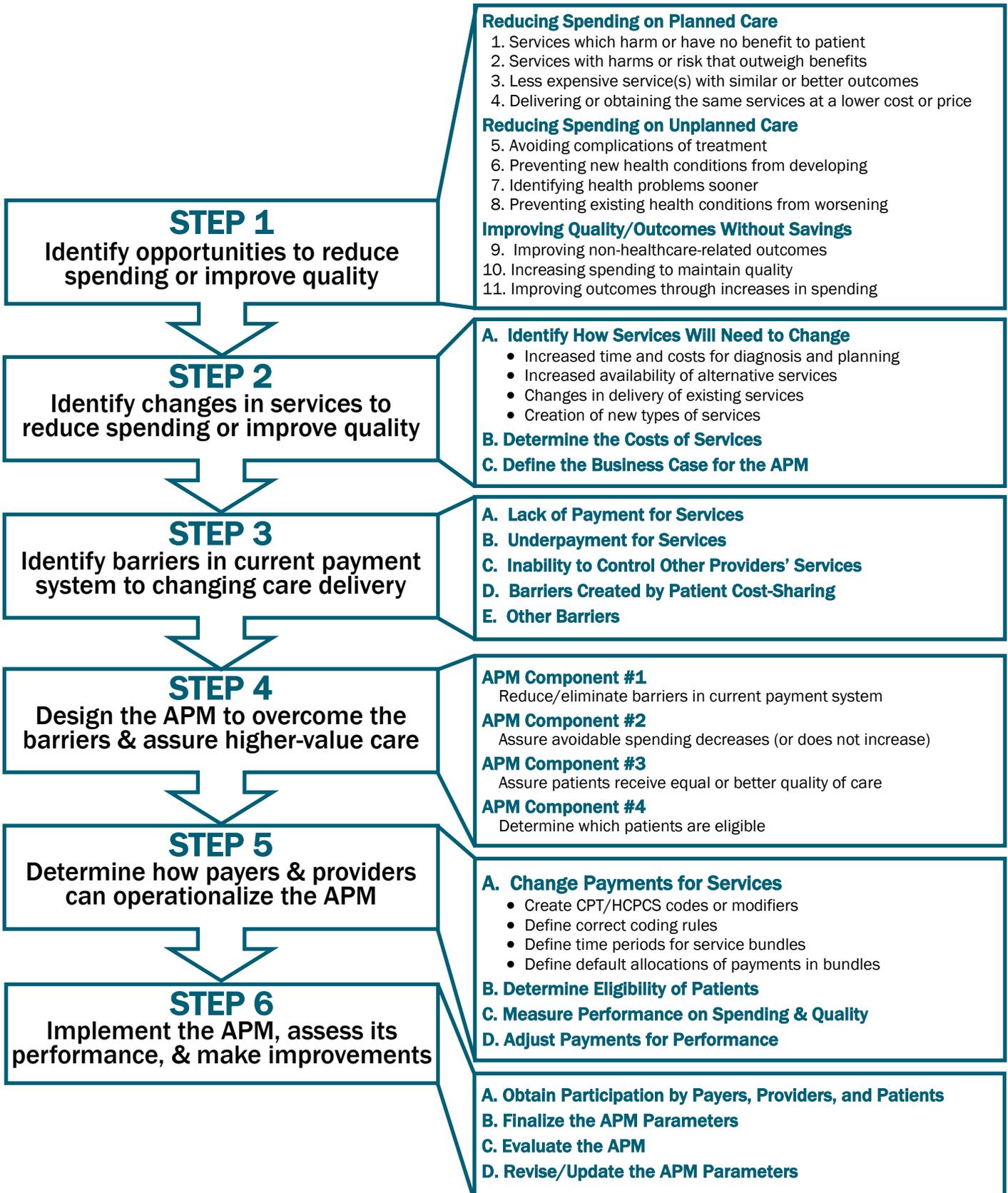
- Step 1:** Identify one or more opportunities for reducing spending and/or improving the quality of care;
- Step 2:** Identify changes in care delivery that will reduce spending or improve quality in those opportunity areas;
- Step 3:** Identify the barriers in the current payment system that prevent or impede implementing the improved approach to care delivery;
- Step 4:** Design the Alternative Payment Model so that it will overcome the barriers in the current payment system and assure the delivery of higher-value care;
- Step 5:** Determine how payers and providers can operationalize the APM as easily and quickly as possible; and
- Step 6:** Implement the APM, assess its performance, and make improvements as needed.

Most current APMs have not been designed to focus on specific opportunities for reducing avoidable spending. Defining the goal of the APM as “reducing the total cost of care” may seem ideal from the perspective of a payer, but it can be highly problematic for both healthcare providers and patients because:

- There are many ways total spending might be reduced that would be harmful for patients. An APM that targets specific opportunities to reduce spending by improving the quality of care will be much safer for patients than an APM that rewards providers for *any* reduction in healthcare spending.
- Providing adequate payments requires knowing what high-value services will need to be delivered to reduce spending or improve quality.
- No individual physician, hospital, or other provider delivers all of the services any individual patient receives or all of the factors affecting the total cost of care for their patients. Accountability needs to be focused on the specific aspects of spending and quality that providers can control.

Consequently, the starting point in creating an APM is to identify specific opportunities for improving outcomes and/or reducing potentially avoidable spending. The APM can then be designed to pay adequately for the necessary services and to hold providers accountable for achieving the expected results.

STEPS TO CREATE AN ALTERNATIVE PAYMENT MODEL



STEP 1: IDENTIFY OPPORTUNITIES FOR SAVINGS & QUALITY IMPROVEMENT

A successful Alternative Payment Model will achieve reductions in healthcare spending in ways that maintain, and ideally improve, the quality of care for patients. Opportunities for doing this can be divided into the following eight categories:

Reducing Spending on Planned Care:

1. Avoiding the use of services that harm or have no benefit for the patient;
2. Avoiding the use of services with harms or risks that outweigh the benefits;
3. Using a different service or combination of services that is less expensive but achieves similar or better outcomes; and
4. Delivering or obtaining the same services at a lower cost or price.

Reducing Spending on Unplanned Care:

5. Avoiding complications of treatment;
6. Preventing new health conditions from developing;
7. Identifying treatable conditions before they worsen; and
8. Preventing existing health conditions from worsening.

“Reducing spending” includes avoiding increases in spending that would otherwise have occurred if utilization of avoidable services is expected to increase in the absence of the APM.

There may also be opportunities to improve the quality of care or outcomes for patients that do not result in any healthcare savings. If there is no change in spending, but quality or outcomes improve, that could still qualify as an APM. If an opportunity for improving quality would require an increase in spending, it would need to be combined with an opportunity for reducing spending in order to be part of an APM.

There are also situations in which spending may need to increase simply to sustain current levels of quality and outcomes, such as addressing the problems of underpayment for services faced by many rural hospitals and physician practices. Since these changes would not qualify as an APM, they would need to be pursued through other types of payment reform.

STEP 2: IDENTIFY CHANGES IN SERVICES NEEDED TO IMPROVE CARE

The existence of an opportunity for reducing a particular aspect of spending while maintaining or improving quality does not automatically mean that savings in that area can be reliably achieved; there must be a systematic way of delivering care differently that can successfully address that opportunity, and any additional spending involved must be less than the savings that are achievable. To determine whether an Alternative Payment Model is feasible, three separate steps are needed:

- **Identify one or more changes to care delivery that are expected to achieve the desired savings or improvement in quality.** An APM is unlikely to be successful unless it is clear there is at least one way to deliver healthcare services differently that can achieve the desired results in terms of savings and quality. The specific ways in which services will need to change must be identified in order to ensure that the APM design adequately supports an improved approach to care delivery.
- **Determine the costs of delivering services under the revised approach to care.** The cost of delivering a service may be very different from what Medicare or other payers currently pay for the service (if they pay for it at all). Even when the goal of the APM is to avoid unnecessary or harmful services, providers may need to spend more time or incur more costs in order to make the decisions to change services or to deliver alternative services. If current payment amounts are less than the costs of delivering desirable services, it may be impossible to sustain those services under

the APM; if payment amounts are higher than costs, reducing payments could provide an additional way to generate savings.

Cost-to-charge ratios cannot be used to accurately determine the true costs of individual services. Moreover, if the volume of services changes under the APM, the cost of delivering services will also likely change. Because a significant proportion of most healthcare providers' costs are fixed, the *average* cost per service will *increase* when fewer services are provided. Consequently, it is not enough to have a cost accounting system that reports what it currently costs to deliver a service; a *cost model* is needed that identifies the fixed costs, semi-variable costs, and variable costs associated with the service and estimates how those costs will change when there are changes in the number or types of services delivered.

- **Determine whether there is a business case for pursuing development of an APM.** If the estimated increase in cost associated with the change in service delivery is less than the savings expected to result from reducing the avoidable spending, the APM can be successful. If not, a different approach to service delivery will be needed that has a lower cost or a larger impact on avoidable spending, or a payment reform other than an APM may be more appropriate.

STEP 3: IDENTIFY THE BARRIERS IN THE CURRENT PAYMENT SYSTEM

The current fee-for-service payment system may create barriers to the delivery of the new or modified services needed to achieve savings and/or improve quality. If the APM does not identify and remove these barriers, it will be unlikely to achieve the desired results. Common barriers to implementing changes in care delivery include:

- **No payment for one or more of the services providers would need to deliver.** The current fee-for-service payment system defines specific payment amounts for over 15,000 different services. Despite this, many payers do not pay at all for a variety of high-value services, such as communications between physicians and patients, communications between primary care physicians and specialists, palliative care services for patients with advanced illnesses who do not qualify for hospice care, and many others. In some cases, payments may only be available for the service in certain circumstances that do not include the patients or providers targeted by the APM.
- **Current payments for the services to be delivered are less than needed to cover the costs of delivering the services.** For example:
 - ◆ **Underpayment for specific phases of care.** The amount of payment may be too low for a service when it is delivered during certain phases of the care process.
 - ◆ **Underpayment for specific kinds of patients.** If there is only one payment amount for delivery of a service, but the amount of time, staffing, or materials required to deliver the service varies significantly from patient to patient, then the provider will be financially penalized for treating the higher-cost patients.
 - ◆ **Underpayment related to volume.** The payment amount may be too low for providers who deliver the service less frequently than others. Because a significant portion of the costs of many healthcare services is fixed, a healthcare provider that reduces the volume of services delivered can experience losses when paid an amount that would be ade-

quate for higher-volume providers. Providers in rural areas will often have higher costs to deliver a service than providers in more densely-populated areas simply because of the lower number of eligible patients.

- ◆ **Underpayment for new services.** There will often be significant startup costs associated with a new service, or a period of time in which costs have to be incurred before revenue can be generated. A payment amount that is adequate to cover ongoing costs may not be enough to enable recovery of startup costs.
- **Healthcare providers are unable to control the types or costs of services delivered by the other providers they rely on for a portion of their patients' care.** Under current fee-for-service payment systems, each provider is paid separately for the services they deliver, and so a provider participating in the APM may be unable to control whether other providers deliver an undesirable service, fail to deliver a service that patients need, or use an unnecessarily expensive method of delivering a needed service.
- **Patients are unable to afford to pay for the services or to pay their share of the cost of services under their insurance plan.** If the patient feels the cost-sharing amount is unaffordable or is not commensurate with the benefit of the service to them, the patient may not seek out or accept a service, even if doing so would enable the insurer to achieve savings on its share of the payments or enable the provider to achieve better outcomes for the patient.

There may also be barriers to delivering the desired services or reducing the avoidable services that have nothing to do with the payment system, such as fear of being sued if a test or service was not delivered, inability to deliver a particular service because of the scope of practice laws in the state, or restrictions in federal and state fraud and abuse statutes. These barriers cannot be addressed by changes in the payment system alone.

STEP 4: DESIGN THE ALTERNATIVE PAYMENT MODEL

An Alternative Payment Model needs four distinct, but interrelated components:

APM Component #1:

A mechanism for reducing or eliminating the barriers in the current payment system that impede delivering the services that would reduce specific types of avoidable spending;

APM Component #2:

A mechanism for assuring patients and payers that the avoidable spending targeted by the APM will decrease (if the goal of the APM is to achieve savings), or that spending will not increase (if the goal of the APM is to improve quality);

APM Component #3:

A mechanism for assuring that patients will receive equal or better quality of care and outcomes as they would with the kind of care delivery they receive under the current payment system; and

APM Component #4:

A mechanism for determining which patients will be eligible for the services supported by the APM.

There are multiple ways to implement each of these components.

APM Component #1: Removing the Barriers in the Current Payment System

If the current payment system creates barriers to delivering the services needed to achieve reductions in avoidable spending, the APM needs to remove those barriers or at least reduce them. The mechanism used to do that depends on the nature of the barriers and on the ways care may be delivered once the barriers are removed. There are at least fourteen options for doing this. These options are not mutually exclusive, and two or more options may need to be combined, either to address multiple barriers in the current payment system or to avoid creating a new type of barrier by using an overly narrowly defined payment change.

Paying for Unpaid Services

Option 1: Pay a Fee When the Service is Delivered. If the barrier to delivering a high-value service is that there is no payment for that service, the most straightforward solution is to simply create a fee for the service. If there are only specific circumstances in which delivery of the service is desirable, those circumstances can be defined as conditions required in order for the fee to be paid. Concerns about potential overuse of the service can be addressed through Component #2 of the APM.

Option 2: Pay for the Service Through a Bundled Fee for a Group of Services. An alternative to paying a separate fee for an individual service is to include the service as a part of a group of services and pay a single “bundled” fee for the group. This can be desirable if the service should always or almost always be delivered together with the other services in the group, if the service is intended as an alternative to one or more of the other services in the group, or if there are different ways of delivering the service itself to achieve the same results. However, bundled payments are not always better, particularly when different patients will need more or fewer of the services in the bundle.

Aligning Payments With the Costs of Services

Option 3: Increase the Payment to Cover Costs. If the payment amount for a service is lower than the cost of delivering that service in most or all circumstances, an obvious solution is to increase the amount of payment to match the cost of delivering the service. If the payment is too low in specific circumstances, then it may be preferable to define a different payment for the service in those circumstances, using Option 1.

Option 4: Stratify Payments by Phase of Care. If there are situations in which the “same” service or group of services is costlier to deliver in one phase of care than another, e.g., when a chronic condition is first diagnosed and treated, payments can be “stratified” by phase, i.e., the amount of payment is determined by both the type of service and the phase of care in which it is delivered.

Option 5: Stratify Payments by Patient Characteristics. If it takes longer to deliver a service to patients with specific characteristics, or if the costs for materials or devices are higher for certain types of patients, higher payments can be defined for the service when it is delivered to pa-

tients with those characteristics. Stratification is usually preferable to “risk-adjusting” payment amounts because of weaknesses in the methodologies used for risk adjustment.

Option 6: Condition-Based Payments. If the cost of delivering a service depends more on the number and types of patients being treated than on the number of times the service is delivered, a “condition-based payment” – paying based on the number of patients treated for a particular condition – will be preferable to paying fees for each individual service. A “condition” could include multiple diseases that require coordinated treatment, and condition-based payments can also be stratified and/or bundled. There will need to be an objective way of defining and documenting the presence of the condition that will trigger the payment.

Option 7: Standby Capacity Payments. There are a number of important healthcare services, such as hospital emergency departments, which must be available in a community regardless of how many patients are treated or whether any patients are treated at all. Fee for service payment is not an appropriate way to pay for these “standby” services, because the services provide a benefit not just to patients who actually use them, but also to the individuals who could have potentially needed them. Standby capacity payments represent a way to ask “potential patients” to pay for the fixed costs of this standby capacity.

Option 8: Volume-Based Adjustments. An alternative approach when services have significant fixed costs is to pay on a per-service basis, but explicitly adjust the payment amount based on the total volume of the services delivered by the provider.

Option 9: Outlier Payments. If there are *individual* patients who have unique characteristics that make the cost of delivering services dramatically higher than average, a provider could receive an outlier payment to cover all or part of the extra costs involved in delivering services to those patients.

Option 10: Cost-Based Payments. A cost-based payment explicitly ties the payment amount to the *actual* cost a provider incurs for delivering a service or combination of services to the specific patients who received the services.

Option 11: Using Multi-Component Payment Structures. Options 1-9 are each designed to align payment with one aspect of costs – either fixed costs, semi-variable costs, or variable costs – but not with all three. Since most services involve a combination of fixed costs, semi-variable costs, and truly variable costs, none of the options is ideal for matching payment to costs at different volumes of services. To address this, a payment model can be created that explicitly includes separate components using two or more options from Options 1-9.

Enabling Control of Services Delivered by Other Providers

Option 12: Multi-Provider Bundled Payment. In a multi-provider bundled payment, a single payment supports all of the individual services delivered by all of the providers who need to work as a team, so all of the included providers can hold each other accountable for what they are doing and how they are paid. Such bundles work best when the providers have agreed to work together as a team and the patient has agreed to use the members of that team for the services included in the bundle.

Modifying Cost-Sharing

Option 13: Modify standard cost-sharing rules. In most insurance plans, the amount that a patient is expected to pay for a healthcare service is determined using some combination of copayments, co-insurance, and deductibles. Special cost-sharing requirements could be created for services delivered under the APM to ensure that they do not discourage the use of desirable services or encourage the use of undesirable services.

Option 14: Create or change last-dollar cost-sharing amounts. Typical cost-sharing requirements are “first dollar,” i.e., the amount that the patient pays is determined first, and then the payer pays the rest. An alternative is to require the patient to pay the “last dollar” of the cost, i.e., if there are two different choices of services or providers, the patient’s cost sharing would be based on the difference in the cost.

COMPONENT #1: Removing the Barriers in the Current Payment System		
Payment Option	Payment Barrier(s) Addressed	Challenges/Weaknesses
1. Pay a fee for the service	No payment for a high-value service	Can encourage unnecessary use
2. Bundled payment for a group of services	No payment for a service that complements or substitutes for other services	Can limit flexibility if patients need different combinations of services
3. Higher payment for the service	Payment is usually below cost	Can encourage unnecessary use
4. Payment stratified by phase of care	Payment too low in some phases	Requires clear definition of phases
5. Payment stratified by patient characteristics	Higher cost of delivering service to certain types of patients	Requires objective way of assessing presence of characteristics
6. Condition-based payment	Cost depends more on number and type of patients than # of services	Can encourage over-diagnosis of condition
7. Standby capacity payment	Service needs to be available even if no patients need or use it	Requires determining minimum capacity needed for service
8. Volume-based payment adjustment	Higher cost for low-volume providers	Can encourage delivery of low volumes of service
9. Outlier payment	Higher cost for specific patients	Can reward inefficiency
10. Cost-based payment	Costs differ for different providers	Can encourage inefficiency
11. Multi-component payment	Cost of services depends on multiple factors	Increases the complexity of payment
12. Multi-provider bundled payment	Multiple providers need to deliver services in a coordinated way	Requires designating a payment recipient and allocation method
13. Modified first dollar cost-sharing	Co-pays, co-insurance, deductibles discourage use of high-value service	Lower cost-sharing can encourage unnecessary use
14. Last-dollar cost-sharing	Different providers/services have similar benefits but different costs	Can discourage use of higher-cost services that have better outcomes

APM Component #2: Creating Accountability for Spending

If the changes in payment included in Component #1 eliminate or adequately mitigate the payment barriers identified in Step 3, then it should be feasible for patients to receive the kinds of services defined in Step 2. However, in order to make these changes in payment, a payer or patient will also want assurance that the expected savings will actually materialize. An accountability component for spending has four distinct elements:

1. One or more *measures of spending or utilization* that the participants in the APM will be accountable for reducing or controlling;
2. A *Target* for each of these measures, i.e., the level that must be achieved or maintained or the change

that must occur in order for the APM to be deemed successful in achieving its goal;

3. A *performance assessment methodology*, i.e., the calculations that will be made to determine whether a specific entity participating in the APM has achieved or maintained the targets.
4. A *mechanism for adjusting payments based on performance*, i.e., what changes will be made in payments if the targets are not achieved.

It is often desirable to have multiple accountability components for different aspects of spending.

1. Defining the Accountability Measures

The APM needs to define the specific aspects of utilization or spending for which the participant in the APM will be accountable and how they will be measured.

If the APM is explicitly intended to reduce or control spending on certain types of services, then the APM needs specific measures for each of those services or the aspects of spending that are to be reduced. This could include:

- Planned reductions in utilization or spending on services delivered by the APM participants.
- Planned reductions in utilization or spending on services ordered from other providers.
- Reductions in utilization or spending on unplanned services that the APM is intended to achieve.
- Spending on complications of treatment related to the new or expanded services under the APM.
- Spending on complications of undertreatment when fewer or different services are being delivered under the APM.
- Spending on substitutions of other services for the services reduced by the APM.
- Spending from increased utilization of a lower-priced service.

Using a “total cost of care” measure may seem simpler and more reliable than defining and measuring spending for specific types of services, but such a measure can be problematic because individual providers generally cannot control all aspects of utilization and spending. Using measures of total spending can also be problematic for the patients who are receiving services supported by the APM because it creates financial incentives for providers to inappropriately delay or withhold needed services. Moreover, the random variation in utilization and spending in a total cost of care measure can hide meaningful reductions in spending that are achieved in specific types of services. These problems can be reduced by using a more narrowly-defined composite measure that includes only services related to the specific condition for which the patient is being treated or to a specific procedure the patient has received (e.g., an “episode

spending” measure). However, use of any kind of composite measure makes it more difficult for providers, payers, and patients to determine whether the APM is achieving savings in desirable or undesirable ways.

In most cases, the best approach will be to use a combination of both service-specific measures and composite measures based on the types of impacts on spending the APM could have. Two or three separate groups of measures or composites could be defined as follows:

- a. **Potentially Avoidable Spending**, i.e., one or more service-specific measures for aspects of spending where the APM is intended to achieve savings. For each of these measures, specific goals for savings would be defined.
- b. **Related Spending**, i.e., service-specific measures, or a single composite measure, focused on specific types of services and spending where increases caused by the APM are possible but undesirable. Here, the goal would be no increase in utilization or spending on these measures of related spending (or an increase smaller than the savings on targeted spending).
- c. **Unrelated Spending**. If there is concern that utilization or spending could increase in other, unidentified areas, an additional broad composite measure of spending could be defined by taking an episode spending measure or total cost of care measure and subtracting the aspects of utilization or spending defined in the first two groups, and monitoring this measure for significant changes.

Instead of measuring spending, it may be preferable to measure *utilization* or *resource use* in order to separate the effects of individual providers’ decisions about which services to use from decisions made by pharmaceutical companies, device manufacturers, and large health systems about the prices they charge for services. Moreover, it will generally be desirable to *stratify* or risk-adjust measures of utilization and spending for differences in patient needs. In addition to defining the types of services for which utilization, spending, resource use, or appropriateness will be measured, a decision must also be made about the *timeframe* in which those services must occur in order to be included in the measure.

2. Setting the Performance Targets for Utilization and Spending

Alternative Ways of Setting Targets

A Target for each of the measures is needed that defines the level of spending or utilization that must be achieved to assure that the business case for the APM is being fulfilled. Two different types of Targets can be defined based on the types of savings that are expected:

- a. **Patient-Level Targets.** Ideally, an APM will define the Target for a measure in terms of the level of service utilization or spending that is appropriate for *each individual patient*, based on that patient's needs. This is easiest to accomplish for an APM that makes changes in *planned* services that are expected to achieve net savings for every participating patient with particular characteristics (e.g., use of a less expensive but equally effective service). When different types or amounts of services are appropriate for different patients, the Target could be defined as adherence to evidence-based clinical guidelines or "pathways."
- b. **Population-Level Targets.** An alternative is to define Targets in terms of the level of utilization or spending to be achieved for a *group of patients*. This is most appropriate in APMs that are designed to reduce *unplanned* services, since any individual patient might or might not have experienced an unplanned service (e.g., a complication of surgery) even without the APM. There are three different ways to set Population-Based Targets:
 - i. **Benchmark-Based Target.** Because spending under an APM is required to be equal to or lower than it would have been in the absence of the APM, most Population-Level Targets, at least initially, will likely be defined as a Benchmark-Based Target using two separate components:
 - a *Benchmark* that defines what level of spending/utilization for the group of patients receiving services supported by the APM is viewed as reflecting "no impact of the APM"; and
 - a *Target Change*, i.e., the minimum or maximum amount by which actual spending or utilization under the APM should differ from the Benchmark.
 - ii. **Evidence-Based Target.** If there is evidence indicating that a specific level of utilization or spending can be achieved that is lower than the level currently being achieved by most providers, then that level of utilization and spending could be set as an Evidence-Based Target, thereby avoiding the need to define Benchmarks and Target Changes.
 - iii. **Competitive Target.** In situations in which there are multiple providers offering services under an APM, the Target could be set through a competitive process.

Alternative Ways of Defining Population-Level Benchmarks

If a Benchmark-Based Target is going to be utilized, three basic methods can be used to define the Benchmark:

- **Prior Performance Benchmark.** This is based on the actual level of spending or utilization during a previous period of time, either for the same patients or for the patients the same provider has treated or managed in the past.
- **Comparison Group Benchmark.** This is based on the actual level of spending for a group of patients who are not participating in the APM but who are similar to those who are in the APM.
- **Counterfactual Benchmark.** This is based on an estimate of what the spending or utilization in the current year would be for the specific patients who are receiving services supported by the APM.

Alternative Ways of Defining Target Changes

Since the Benchmark for a measure is intended to represent the level of spending/utilization that reflects "no impact" of the APM, the Target Change must define the magnitude of the desired impact of the APM. There are four different approaches that could be used to define the Target Change:

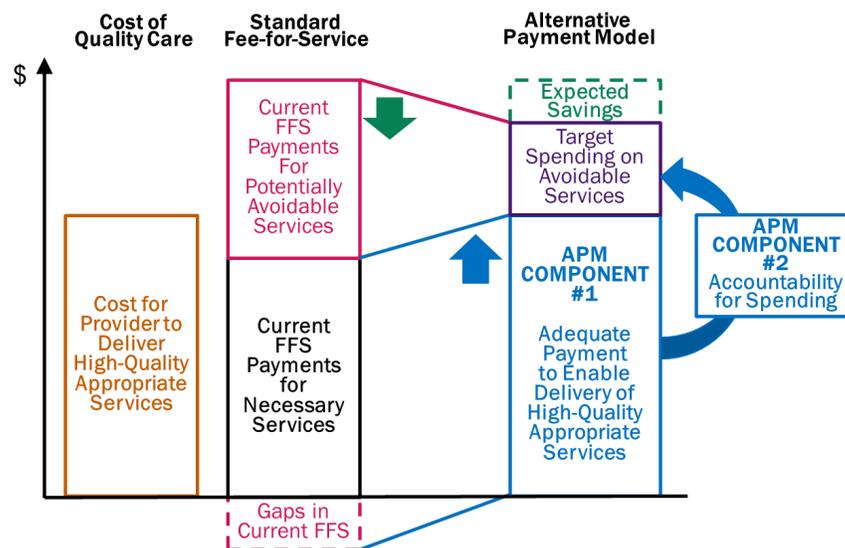
- **Minimum/Maximum Change Needed for Success.** If the APM is intended to reduce utilization or spending, the Target Change could be set at a level that achieves sufficient savings to offset any expected increases in spending on desirable services. If the goal is to avoid an increase in spending, the Target Change could be defined as either zero or an increase that would be less than the net savings expected for other services under the APM.
- **Change Achieved by a Comparison Group.** Since there is frequently uncertainty regarding whether unplanned care will occur and the extent to which changes in planned care will be able to affect it, the Target Change could be defined based on what other participants in the APM have achieved, or what participants in other initiatives have achieved.
- **Statistically Significant Change.** Since there is a considerable amount of patient-to-patient variation in utilization and spending on services, and not all of this variation is controllable by the APM participant or even predictable, the Target Change could be defined in such a way as to provide confidence that the change was not due to random variation.
- **Desired Level of Change.** The Target Change amount could also be set at a level that would achieve a specific amount of savings or a specific level of utilization that is desired by the payer and/or the providers and is believed to be achievable.

Issues in Defining Spending/Utilization Targets

Several additional issues need to be addressed in setting Targets for spending or utilization:

- **Prospective vs. Retrospective Targets.** A *Prospective Target* is determined before the beginning of the time period in which performance is going to be evaluated, whereas a *Retrospective Target* is determined afterwards. In general, it is preferable to use Prospective Targets so that providers know what is required for success and payers and patients can predict how much they will need to spend.
- **Common Targets or Participant-Specific Targets.** Although it is easier for a payer to assure that overall savings are being achieved if each provider participating in an APM is required to generate savings, this can penalize providers who had already found ways to reduce avoidable spending prior to the APM, and it can result in individual patients and payers paying more for care from APM participants that were able to “achieve savings” simply by partially reducing use of services they had been overutilizing in the past.

- **Ensuring Similarity of Patients in Calculating Benchmarks.** If the patients used in calculating benchmarks are different from the patients participating in the APM, failure to adjust for the differences could result in the provider being inappropriately rewarded or penalized. Making adjustments solely based on diagnosis codes can be problematic, both because many important differences in patients are not captured by diagnosis codes and because the completeness and accuracy of coding is likely to be higher for the patients in the APM.
- **Revising Targets and Changing the Target Methodology Over Time.** Changes in costs, technology, and medical evidence require that Benchmarks and Targets be updated regularly. In addition, it may be necessary to change the methodology for setting Benchmarks or to move to a different approach to setting Targets if there is no longer a good basis for defining comparison groups.



3. Assessing Performance on Utilization and Spending

An assessment methodology is needed to determine the extent to which any difference between the measure and the Target was due to the APM participant’s performance rather than errors in calculation or measurement or the effects of uncontrollable factors, rare events, or random variation. Because there is a large amount of unexplained variation in most measures of utilization and spending, there will be considerable uncertainty as to whether a difference between the measured level of utilization/spending and the Target represents an actual change in utilization/spending and whether the change is attributable to actions by the providers participating in the APM.

Although it is important to recognize the impacts of random variation and to try to avoid drawing incorrect conclusions because of it, an excessive focus on statistical significance can be problematic. Although requiring high levels of statistical significance theoretically reduces the chance of inappropriately determining that an APM has been successful, it also increases the chance of inappropriately determining that an APM has failed (i.e., reducing “Type I errors” increases “Type II errors”). These tradeoffs are particularly important to consider if only a small number of patients are participating in the APM, if the patients are diverse, and if the performance period is short. A good performance assessment methodology should consider *both* the magnitude and the certainty of a provider’s performance in determining success or failure.

4. Making Performance-Based Adjustments to Payments

Once the spending measures, targets, and methods of assessing performance are defined, the final step is defining the mechanism of accountability, i.e., the actions that will be taken if the actual performance on one or more of the measures is determined to have fallen short of the target level. There are five basic options for accountability:

Option 1: Penalties or Bonuses in Addition to Service-Based Payments

Under this option, the healthcare provider that is participating in the APM is paid for delivering the desirable services using whatever methodology is defined in Component #1, but the provider is required to pay a penalty if the Targets on one or more utilization/spending measures are not achieved.

In general, it is desirable to make the penalty proportion to performance on the measure. In addition, a method is needed for determining the absolute amount of the penalty. Two options for determining the absolute amount of the penalty are:

- Basing the penalty on the amount the provider is paid for *planned* services.
- Basing the penalty on the *Target Change* in spending for the provider's patients.

In addition, limits can be placed on the penalties in order to limit the financial impact of poor performance on the provider, and bonuses can be used in addition to penalties to reward and encourage performance that is better than the Targets.

Option 2: Outcome-Based Payments for Services

Under this option, the APM participant would receive no payment under the APM for an individual patient unless (a) the provider delivered the services the patient was supposed to receive, and (b) that patient did *not* receive the planned or unplanned services that the APM was supposed to avoid.

From the patient's perspective, this is how a value-based payment should work: a patient only pays for services (or only pays the full price) if they received the right services and those services achieved the desired outcome. Under Option 2, the amount the APM pays for planned services would need to be increased to reflect not only the cost of the services but the likelihood that

the provider will achieve the Target. Limits could also be placed on the maximum amount that a provider could lose.

Option 3: Bundled/Warrantied Payments for Services

Under Option 2, a patient or payer would not have to pay a provider for planned services if the services did not achieve the desired outcome, but the patient/payer would still have to pay for the unplanned services or increases in spending they had expected to avoid. Under Option 3, the provider would be expected to use the payment not only to support the planned services but also to pay for any unplanned services that were supposed to be avoided.

This is analogous to a warranty on a product or service. The APM participant is not guaranteeing that no complications or other unplanned services will occur, it is merely agreeing to pay to treat them if they do occur without receiving any additional payments from the patient or payer. Similar to warranties in other industries, the amount of a bundled/warrantied payment for a service would be higher than payments today because it would cover unplanned services that would otherwise be paid separately.

Option 4: Terminating a Provider's Participation in the APM

Options 1-3 all assume that a provider that fails to meet a Target will pay some type of financial penalty and continue participating in the APM (if they wish to do so). A fourth option is to simply terminate the provider's participation in the APM altogether if the provider does not achieve success on the performance measures. This allows greater flexibility to consider the circumstances that may have led to failure or success in meeting the targets.

Option 5: Terminating the APM

A final option is simply to stop using the APM altogether. If APM participants collectively are not succeeding in reducing spending or maintaining spending while improving quality, then it makes sense to modify the design of the APM or to terminate it and develop something different.

APM Component #3: Creating Accountability for Quality

It is not enough for an Alternative Payment Model to maintain or reduce spending; there must also be a way of assuring that the quality of care for patients is maintained or improved. There are four distinct elements in an accountability component for quality:

1. One or more *measures of quality* that need to be maintained or improved by the services supported by the APM;
2. *Targets* for the level of quality that must be maintained or the improvement that must be achieved in

each aspect of quality in order for the APM to be deemed successful in achieving its goal;

3. A *performance assessment methodology* to determine whether a specific provider participating in the APM has achieved the quality Targets; and
4. A *mechanism for adjusting payments based on performance*, i.e., what changes will be made in payments if the Targets are not achieved.

1. Defining the Accountability Measures

Many current APMs have chosen to hold APM participants accountable only for aspects of quality where measures already exist. However, if those measures do not match the specific aspects of quality likely to be affected by the APM, they will not provide adequate protection for patients and they will divert providers' attention from the intended goals of the APM. Determining whether existing or new measures are most appropriate requires three separate steps:

- a. Identifying the aspects of quality affected by the APM;
- b. Determining how to assess changes in quality; and
- c. Determining whether and how data needed to make such assessments can be obtained.

a. Identifying the Aspects of Quality Where Accountability is Needed

There are four general areas that should be examined to determine what quality measures are needed:

- Aspects of quality where the APM is intended to make improvements.
- Aspects of quality that could be harmed by changes in services that are explicitly encouraged by the APM.
- Aspects of quality that could be harmed by incentives created through the payment methodology or spending accountability components of the APM.
- Aspects of quality necessary to ensure accurate payment under the APM (e.g., accuracy of data on diagnosis and outcomes).

b. Determining How to Assess a Particular Aspect of Quality

Ideally, the quality of care would be assessed based on the outcomes achieved for patients. However, relatively few outcome measures have been developed and even fewer are currently in use because of the challenges in collecting and interpreting outcome measures. In addition, most outcomes are not totally under the control of healthcare providers. "Process" measures, i.e., measures of whether a particular activity was performed,

are more commonly used because they are easier to collect and because they tend to focus on aspects of care delivery that the provider can control. However, process measures can be problematic if a goal of the APM is to enable care to be delivered in different ways. A third option is "intermediate outcomes," such as laboratory test results and other biomarkers, if they are highly correlated with longer-term outcomes.

The choice of measures should be based on the goals of the APM and the care it is designed to support:

- **Outcome measures** will be preferable when providers can control the factors that affect outcomes.
- **Process measures** will be appropriate when the goal is to achieve more reliable or efficient delivery of current evidence-based processes.
- **A combination of process and outcome measures** will be desirable when the goal is to deliver care in ways that are not supported by the current payment system. The process measures would ensure that desirable changes are made in care, and the outcome measures would ensure the changes are having positive impacts on the patients.

c. Obtaining Data to Assess the Quality of Care

No matter which quality measures would be most desirable in theory, it will only be possible to use measures for which the necessary data can be obtained in an accurate, reliable, affordable, and timely way. If data that match the definition of quality needed for the APM are not collected currently, new or modified data will be needed, and the APM will need to pay enough to cover the costs associated with collecting these data.

For each quality measure, the Target level of quality the APM participants will be expected to achieve must be defined. At a minimum, the Target should ensure that the quality of care did not decrease, and if the APM is intended to improve quality, the Target would need to reflect that.

2. Setting the Performance Targets for Quality

Using Patient-Level Targets to Ensure Quality Does Not Decrease

Most current quality measures cannot be used to ensure that the quality of care is not harmed by an APM. These population-based quality measures calculate the percentage of patients for whom a process was performed or a particular outcome level was achieved, and compare that percentage to a previous period or to patients who are not participating in the APM. However, the fact that a similar or higher *percentage* of patients is receiving high quality care under the APM does not mean that every patient is receiving equal or better quality care.

From a patient's perspective, what matters is whether the APM is maintaining or improving the quality of care that *individual patient* receives, not what happens to *other* patients. Consequently, the starting point in setting quality targets for an APM is to define appropriate *Patient-Level Targets*, i.e., the threshold(s) that will be used for determining if an *individual* patient is benefiting or being harmed by participating in the APM.

There are several approaches that can be used to define Patient-Level Targets for quality:

- **Maintaining Prior Levels of Quality**, if the patients have been receiving treatment for the same condition in the past;
- **Achieving Evidence-Based Standards or Guidelines**;
- **Achieving Statistically Significant Improvement**;
- **Achieving Clinically Important Improvement**; and
- **Achieving Patient-Specific Goals**.

Using Population-Level Targets to Assess Improvements in Quality

If the APM is intended to *improve* quality on a particular measure, a *Population-Level Target* can be used in addition to a Patient-Level Target. Although it is problematic if any individual patients are being harmed by participation in the model, it is not necessary that every patient receive better care in order for the APM to be deemed successful in improving care, just as an APM can be successful financially if savings are achieved for some but not all patients. Consequently, if the APM is expected to improve quality, two sets of Targets should be defined:

- a Patient-Level Target that defines the *minimum* level of quality that must be achieved for *each* patient; and
- a second Patient-Level Target that defines the higher-than-minimum level of quality that is desired for each patient, and an associated Population-Level Target defining the proportion of patients who need to achieve the higher Patient-Level Target in order for the APM to be viewed as successful.

There are three basic approaches that can be used to define Population-Level Targets for quality:

- **Status Quo-Based Targets**, i.e., improvements in quality compared to current or recent quality levels for the same or similar patients. This requires both a method of defining the "Status Quo" and also defining the Target Change from the Status Quo. Alternative ways of defining the Target Change include:
 - ◆ **Goal-Based Change**, e.g., the level of improvement that would be viewed as sufficient by either payers or providers to justify implementing the APM.
 - ◆ **Statistically Significant Change**, i.e., the minimum change needed to provide assurance that a change is not due to random variation.
 - ◆ **Clinically Important Difference**, i.e., the minimum change needed to be perceived by patients as an improvement in one or more outcomes.
 - ◆ **Comparison Group Change**, i.e., the change in quality for a comparison group not participating in the APM.
- **Evidence-Based Targets**, if there is research showing the quality of care or outcomes that can consistently be achieved for the types of patients participating in the APM when they receive the services the APM is designed to support.
- **Competitive Targets**, i.e., allowing individual providers to determine the level of quality they believe they can achieve.

Issues in Defining Quality Targets

The issues described in Component #2 with respect to targets for utilization/spending measures also apply to the targets for quality measures.

- **Prospective vs. Retrospective Targets**. In general, it is preferable to use Prospective Targets so that providers know what is required for success.
- **Common Targets or Participant-Specific Targets**. Although Participant-Specific Targets can encourage participation by lower-performing providers, they can be problematic from the perspective of patients since they result in the same amount of payment for different levels of quality.
- **Revising Targets and Changing the Target Methodology Over Time**. Changes in technology and medical evidence require that Quality Targets be updated regularly. In addition, it may be necessary to change the methodology for setting Targets if there is no longer a good basis for defining comparison groups.

3. Assessing Performance on Quality

As with Component #2, a methodology is needed to determine the extent to which any difference between the measured level of quality and the Target was due to the APM participant's performance rather than errors in calculation or measurement or the effects of uncontrollable factors, rare events, or random variation. Requiring high levels of statistical significance reduces the chance of inappropriately determining that an APM has been suc-

cessful, but increases the chance of inappropriately determining that an APM has failed (i.e., reducing "Type I errors" increases "Type II errors"). These tradeoffs are particularly important to consider if only a small number of patients are participating in the APM, if the patients are diverse, or if the performance period is short.

4. Making Performance-Based Adjustments to Payments

Once the quality measures, Targets, and methods of assessing performance are defined, the final step is defining the mechanism by which APM participants will be penalized or rewarded based on how actual performance compares to the Targets. There are five options, which are similar, but not identical, to those described in Component #2:

Option 1: Penalties or Bonuses in Addition to Service-Based Payments

Under this option, the healthcare provider that is participating in the APM is paid for delivering the desirable services using whatever methodology is defined in Component #1, but the provider is required to pay a penalty if the Targets on one or more quality measures are not achieved.

In general, it is desirable to make the penalty proportional to performance on the measure. Three different approaches can be used to determine the absolute amount of the penalty:

- Basing the penalty on the *perceived value of quality*, i.e., a dollar amount would be assigned to the shortfall in quality based on the patient's or payer's view of the value of achieving the Target.
- Basing the penalty on the *amount of payment for planned services*, e.g., a percentage of the payment the provider in the APM would have received if the Target had been achieved.
- Basing the penalty on the *penalty or bonus for utilization/spending* in Component #2. This approach is used in many APMs, but it is undesirable because it can result in no penalty for quality problems, regardless of how serious they are, as long as spending targets are met.

Limits can be placed on the penalties in order to limit the financial impact of poor performance on the provider.

It is challenging to provide bonuses for higher-than-expected quality under an APM because the bonus could potentially increase overall spending under the APM.

Option 2: Outcome-Based Payments for Services

Under this option, the APM participant would receive no payment under the APM for an individual patient unless the provider achieved the Patient-Level Targets for that individual patient. From the patient's perspective, this is how a value-based payment should work: a patient only pays for services (or only pays the full price) if they received the right services and those services achieved the desired outcome. However, this approach would work best for quality measures where it is feasible for a provider to achieve nearly 100% success.

Option 3: Warranted Payments for Services

Under Option 2, a patient or payer would not have to pay a provider for planned services if the services did not deliver adequate quality care, but the patient would still experience the negative effects of the poor-quality care. Under Option 3, the provider might still receive the standard payment for the services that were delivered to the patient, but the provider would pay the patient some amount of compensation to offset the impacts of the poor-quality care.

Option 4: Terminating a Provider's Participation in the APM

Options 1-3 all assume that a provider that fails to meet a Target will pay some type of financial penalty and continue participating in the APM (if they wish to do so). A fourth option is to simply terminate the provider's participation in the APM altogether if the provider does not achieve success on the performance measures. This allows greater flexibility to consider the circumstances that may have led to failure or success in meeting the targets.

Option 5: Terminating the APM

A final option is simply to stop using the APM altogether. If APM participants collectively are not succeeding in maintaining or improving the quality of care, then it makes sense to modify the design of the APM or to terminate it and develop something different.

APM Component #4: Defining the Eligible Patients

Even if an Alternative Payment Model is successful in reducing use of unnecessary services, there is the risk that the services supported by the APM will be overutilized in ways that can compromise its success in achieving savings. In order to address this, eligibility criteria can be defined that limit participation to the patients who would have been most likely to receive the unnecessary services and/or to benefit from the services supported by the APM.

However, caution is needed to avoid having eligibility criteria encourage overdiagnosis or overtreatment. Narrowly-defined eligibility criteria can create a perverse incentive for both the patient and the provider to find ways for the patient to meet the criteria in order to receive desirable services available only through the APM. An alternative is to stratify the payment amounts and accountability measures in the APM, so that patients with lower levels of need can still participate but receive services matched to their needs.

It is essential that the determination of whether a patient is eligible for an APM be made *prospectively*, i.e., before the provider participating in the APM begins delivering services supported by the APM to the patient. Many current Alternative Payment Models make the determination of whether a patient is participating in the APM *retrospectively*, i.e., after services have already been delivered, but this approach, and the “attribution”

methodologies used to implement it, creates a number of serious problems that are virtually impossible to overcome.

In most cases, the eligibility determination should be made by the provider(s) of services, not by the payer, particularly if the eligibility criteria are based on patient characteristics that are not currently recorded on standard claims forms. Prospective eligibility determinations also enable the patient to understand what services they can expect to receive and agree to whatever actions they will need to take in order for the providers in the APM to achieve the goals of the APM. A Patient-Provider Care Agreement could be required as part of the eligibility criteria for the APM to ensure that both the provider and patient have discussed and agreed to their mutual responsibilities.

It is also important to ensure that the providers participating in an APM do not selectively avoid patients who need more services and/or are less likely to have favorable outcomes (i.e., “lemon-dropping”) or to limit their services only to the patients who are likely to have the most favorable outcomes (i.e., “cherry-picking”). This can be done by identifying the factors that affect how many services a patient will need and the outcomes they will experience and incorporate those factors into the design of the APM so that providers receive appropriate payments for higher-need and lower-need patients.

Finalizing the APM Design

There are multiple options for designing each of the four components of an Alternative Payment Model. The advantages and disadvantages of the different options will depend on the specific types of opportunities for savings and quality that are being pursued, the approaches to care delivery that will be used to address those opportunities, and the specific barriers in the current payment system that need to be corrected. In addition, the choice of options within each component will also depend on which options are chosen for other components.

It is likely that one of the following four designs will be appropriate in most situations where an APM is needed:

- **Accountable Payment for Service.** A provider receives a new or revised payment for delivering a specific service to patients, and the payment is reduced if targets for spending on specific services and performance on quality measures are not achieved.
- **Accountable Bundled Payment.** A provider or team of providers receives a bundled payment to enable delivery of a group of services to patients or to treat a particular condition, and the payment is reduced if targets for spending on specific services and performance on quality measures are not achieved.

- **Outcome-Based Payment.** A provider is only paid for a service or group of services if standards or targets for quality and spending are achieved.
- **Bundled/Warrantied Payment.** A provider or team of providers receives a bundled payment to deliver a group of services to patients, and the provider team is responsible for using the payment to cover the costs of necessary services and also to pay for avoidable services or services needed for complications of treatment.

Once a preliminary APM design has been developed, analyses should be performed to ensure that the APM design would:

- Remove or adequately mitigate the barriers in the current payment system to enable the desired services to be delivered; and
- Pay amounts for services and achieve levels of savings and quality that create a desirable business case for both payers and providers to implement the APM. This includes ensuring that (a) payments will be adequate to cover the costs providers will incur in delivering services, and (b) the savings expected to be generated will be sufficient to offset any increases in payments compared to the current payment system.

STEP 5: OPERATIONALIZE THE APM DESIGN

Once decisions have been made about the options for each of the components of the Alternative Payment Model, additional details are needed in order to *operationalize* the APM. Mechanisms are needed for making determinations as to whether and how much providers participating in the APM should be paid for specific patients in specific situations, and these mechanisms need to be feasible for payers and providers to implement.

An APM will be easiest to operationalize if it can use existing billing systems, claims payment systems, and data collection mechanisms to the maximum extent possible. Even though current claims forms and coding systems were designed for the current fee-for-service system, the same forms and systems can also be used to operationalize most aspects of APMs by translating the structure of the APM into the “language” of billing and claims payment systems, i.e., procedure codes, modifiers, diagnosis codes, edit processes, etc.

A. Operationalizing New and Different Payments for Services

Most of the options for paying providers differently under Component #1 can be operationalized by adding one or more new codes to the Current Procedural Terminology (CPT®) or Healthcare Common Procedure Coding System (HCPCS) lists. Payers would only pay a provider for one of these new codes if the provider is participating in the APM.

This approach can also be used to pay a higher (or lower) amount for an existing service when it is delivered to a patient who is part of an APM. A new code would be established for the service to easily distinguish when a different amount should be paid. Alternatively, a CPT/HCPCS modifier could be added to the existing code to distinguish when a different amount should be paid. Similarly, if payment amounts are to be stratified by patient characteristics or phases of care, different codes could be established for each stratum or phase.

Although most current CPT/HCPCS codes describe individual services, there are also currently codes that define a bundle of services. Many current codes also have a “global period” that defines the period of time that a payment is supposed to cover. These same approaches to bundled codes and “global periods” can be used to define a bundled payment or condition-based payment under an APM.

If the APM is intended to pay for a new service instead of an existing service, or to pay for a bundle of services instead of the individual services, this can be operationalized through modifications to the files used as part of the National Correct Coding Initiative that define when two codes cannot be billed at the same time.

Multi-provider bundled payments can be operationalized in several ways:

- define a bundled payment code and make the payment for that code to one provider or to an entity rep-

resenting multiple providers that will then divide the payment among the participating providers.

- pay each provider a pre-determined allocation of the bundled payment amount.
- pay each provider a reduced amount for their individual services as the services are delivered, and then pay the remainder after comparing the total payments to the bundled payment amount.

B. Operationalizing Eligibility Determinations

The approach to coding and billing described above can also be used to operationalize prospective eligibility determinations for patients and eliminate the need for problematic retrospective attribution systems. When a provider submits a claim form for a patient using a billing code that is created specifically for the APM, the provider would be explicitly indicating that the patient was eligible for the APM and that the provider agreed to take accountability for achieving spending and quality Targets for that patient as required under the APM. If a patient chose to transfer their care to a different provider, the new provider would bill for the appropriate code, and the payer would know immediately that accountability had shifted to the new provider, rather than waiting for calculations to be made under an attribution methodology.

C. Operationalizing Accountability for Spending and Quality Performance

Operationalizing the accountability components of the APM requires obtaining the data needed to calculate spending and quality measures as well as modifying payments based on a provider’s performance on those measures

1. Measuring Performance

Some of the aspects of utilization and spending for which a provider will be accountable under an APM are services that the provider either delivers or orders. If there are performance measures that focus solely on these services, the provider should be able to calculate the measures in order to determine performance. However, if performance measures include services that are not directly delivered or ordered by the provider (e.g., emergency department visits by a primary care physician’s patients), claims data maintained by the patient’s health insurance plan will be needed to ensure all aspects of utilization and spending are included.

However, using claims data can be problematic for some types of utilization measures (e.g., potentially avoidable services or spending on complications) if the information needed to determine whether a particular service should be included in the measure is not available in the data. Claims have also been the most common source of data for the quality measures that are used in payment systems, and this has been problemat-

ic because key data elements needed to accurately calculate the measure are recorded in electronic health records but not on claims forms.

Many of the weaknesses in claims data can be addressed simply by creating additional CPT/HCPCS codes or modifiers and/or additional ICD-10 diagnosis codes and asking providers to record the codes on claims forms. It will likely be more efficient for providers to extract the information from their EHR and report it using their billing system than to have payers create a quality reporting system that is separate from the billing and claims payment system and then trying to merge the data. If the data needed are not currently being collected, the provider could use whatever method for data collection is most feasible and report the results through standard billing and claims data systems using codes designed for that purpose. For example, patient outcome measures could be collected by surveying patients and then reporting the information using codes recorded on claims forms.

An advantage of using CPT/HCPCS codes for reporting quality measures is that it easily allows a provider that is submitting the code to be paid if there is a significant cost associated with collecting and submitting the data. This would also provide a mechanism for compensating providers who are not participating in the APM for collecting quality and utilization data needed for comparison purposes.

2. Performance-Based Adjustments to Payments

Penalties for failure to achieve a patient-specific performance target (Option 1 in Components 2 and 3) can be operationalized relatively easily by (a) decreasing the standard amount that is paid for the services and then (b) making an additional payment for each patient for whom the performance target is reached. The amount of the reduction in the payment for services would be such that when the provider achieved the minimum performance level needed to avoid a penalty, the sum of the additional payments would be equal to the sum of the

reductions in the payments for the services/conditions, i.e., the provider would receive the same amount of revenue as if there was no performance adjustment. This is equivalent to what is commonly described as a “withhold” – a portion of the provider’s payment is withheld and paid only after the necessary performance has been achieved.

Outcome-Based Payments (Option 2 in Components 2 and 3) can be operationalized by requiring that the relevant Target(s) be achieved before a provider could submit a claim for payment. For outcomes that can only be measured after a long period of time, it may be desirable for the provider to receive a partial payment when the service is delivered, and then the balance of the payment when the outcome is achieved. Two separate CPT/HCPCS codes could be created for this purpose.

Under Bundled/Warranted Payments (Option 3 in Components 2 and 3), if an avoidable service is delivered, or if an additional service is needed to correct a defect in quality (or if some form of compensation were to be paid for the defect), the accountable provider would be required to pay for that from the bundled/warranted payment.

STEP 6: IMPLEMENT THE APM

Four sets of activities are needed for successful implementation of an APM:

- Obtaining agreements by payers, providers, and patients to participate in the APM;
- Finalizing the details of the APM design;
- Evaluating the APM to make decisions about continuation/expansion; and
- Updating the APM parameters over time.

A. Obtaining Participation by Payers, Providers, and Patients

An APM is only a concept until at least one payer agrees to implement it, at least one provider who is paid by that payer agrees to participate, and at least some of the patients insured by the participating payer and receiving care from the participating provider are willing to accept the different approach to care delivery and payment.

1. Encouraging Participation by Payers

Many payers have failed to implement APMs even when there are significant opportunities for savings and there are documented barriers in the current payment system that prevent those opportunities from being achieved. There are several common reasons for this:

- administrative costs for payers to implement the APM;
- disincentives for insurance companies to encourage reductions in healthcare spending;
- benefits to payers of being a “free rider;” and
- barriers in provider contracts.

One or more of the following approaches will likely be needed to encourage payer participation:

- designing the APM to work within existing payer administrative systems.
- using a similar approach to coding as in other APMs.
- designing APMs in ways that can be used with self-insured purchasers.
- requiring payers to publicly disclose the payment methods they use.
- prohibiting provisions of payer-provider contracts that limit the ability to implement desirable APMs.

Purchasers, such as businesses and union trusts that pay for services or buy insurance on behalf of their members, are those who ultimately suffer when spending is higher than necessary, and they can take additional actions to encourage payers to implement APMs:

- selecting payers based on APM participation.
- contracting for insurance and care delivery through purchaser coalitions.
- using direct purchaser-provider contracting.

Providers can also encourage payer participation by:

- refusing to contract with payers who do not implement APMs.
- developing the capability to contract directly with purchasers or to sell insurance products.

2. Encouraging Participation by Providers

Lower-than-expected participation in APMs is often attributed to a preference by providers for traditional fee-for-service. However, in most cases, there are other reasons that providers don't want to participate in APMs, including:

- problems with the design of the APM;
- a small number of payers using the APM;
- the inability to cover extra costs incurred during the transition to the APM;
- lack of reserves to manage financial risk;
- lack of data to estimate potential savings and risks;
- no assurance of stability or continuation of the APM;
- failure of the APM to address specific types of patient needs or unique issues in the community;
- requirements in federal or state laws or regulations that prohibit or limit the ability to implement the APM; or
- unwillingness of the provider to make the reductions in cost or improvements in quality needed to succeed.

APMs are far more likely to be successful if providers are participating willingly. Rather than trying to mandate that providers participate in APMs they find problematic, it makes sense to design the APMs in ways that avoid the problems described above by:

- involving providers in the design of APMs.
- designing APMs using Regional Health Improvement Collaboratives or with state government oversight.
- standardizing designs and measures where possible, but allowing flexibility where necessary.
- enabling providers to access claims data or other sources of information on the services their patients are receiving that are relevant to the APM.
- encouraging payers to participate in Medicare APMs both before and after providers begin participation.
- enabling Medicare to participate in APMs that are being used by private payers.
- reducing the higher financial risks for providers during the initial implementation period for the APM.
- revising laws and regulations that create barriers to implementing APMs.
- refusing to use providers who do not participate in the APM.

3. Encouraging Participation by Patients

The fact that an APM is viewed favorably by payers or providers does not necessarily mean it is desirable from the perspective of the patients who would be receiving healthcare services supported by the APM. A patient will be understandably concerned about an APM if it:

- forces the patient to receive their care from a narrow list of providers that were selected based primarily on the price the providers were willing to charge rather than the quality of care they committed to provide;
- requires the patient to pay more in cost-sharing than they would have paid under the fee-for-service system for the specific services they receive;
- financially penalizes the patient's physician if the physician has to order more services or more expensive services to meet the patient's needs;
- financially rewards a provider if that provider delivers fewer services than the patient needs;
- requires the patient to pay for services even if the quality of care that patient received is poor, as long as the quality of care for most other patients was acceptable; and/or
- fails to evaluate the outcomes achieved or the quality of care delivered for the specific types of health problems the patient has.

At the other extreme, some patients who could potentially benefit from an APM might be unable to do so if the design of the APM would cause providers to lose money caring for those patients. For a patient who has multiple, unusual, or complex needs, the APM should:

- provide higher payments to the provider to cover the costs of the additional time or resources needed to care for that patient;
- exclude or adjust for the legitimately higher utilization or spending on the patient when determining penalties or bonuses for utilization/spending;
- exclude or adjust for differences in care delivery or outcomes when determining penalties or bonuses based on quality.

If APMs are going to be attractive to patients who have choices, they need to be designed to benefit the patients, not just payers and providers. In order for providers to be willing and able to care for patients with higher needs, APMs need to be designed so as to not penalize the provider for taking care of those patients. The solution to both problems is to design an APM to be as *patient-centered* as possible by including the following characteristics:

- setting payment amounts based on patient needs.
- focusing accountability for spending on avoidable services and costs.
- hold providers accountable for quality for each individual patient.

B. Finalizing the APM Parameters

In many cases, it will be difficult to specify the “right” payment amounts and targets for spending and utilization before an APM is actually implemented. Information on costs and achievable performance levels can only be obtained from providers that are delivering services in a different way, but providers cannot deliver services in that way without having an alternative payment model to support them. The more innovative the APM – i.e., the more it differs from the current payment system – the more likely there will be a need for an initial “beta testing” process and potentially additional rounds of refinement after the APM is implemented more widely. The beta testing phase will involve:

- participation by a limited number of interested providers;
- using “best estimate” parameters to initiate APM testing;

- protecting providers, payers, and patients against financial harms during the beta testing process; and
- providing extra resources to enable data collection by providers.

The purpose of beta testing is to *refine* the APM, not to evaluate whether it “works.” In fact, it is likely that an evaluation conducted before an APM has been adequately refined will conclude that the APM is less effective in reducing costs or improving quality than it would ultimately be, and this could cause it to be terminated prematurely or discourage other payers or providers from implementing it.

C. Evaluating the APM

Newly implemented APMs should be evaluated in order to identify and correct any problems. However, an evaluation of an APM must be structured correctly and its results must be interpreted properly. The primary focus of the evaluation should *not* be to determine whether spending was lower and/or quality/outcomes were better, because an APM does not *directly* reduce healthcare spending or improve the quality of care. Instead, the focus should be on whether the APM successfully changed the aspects of payment that were viewed as barriers to delivery of services in a different and better way. If the APM successfully removes the payment barriers it was intended to remove but savings are not achieved or quality is not maintained or improved, the care delivery model may need to be improved, or additional actions besides the change in payment may be needed to support the desired outcomes.

Defining APMs as time-limited demonstration projects can have the perverse effect of reducing the likelihood of success, since healthcare providers are unlikely to funda-

mentally change the way they deliver care in response to a payment change that may only last a few years. Payers and providers should make a commitment to continue implementing an APM for a long enough time to ensure that changes in care delivery can be fully implemented and to recoup the costs incurred in participating in the APM. Payers should also agree to modify the APM in an effort to correct any weaknesses before terminating it.

It is undesirable to mandate participation of providers in an APM simply to support a more robust evaluation. If it is not yet clear that the APM is designed correctly, it is inappropriate to force providers and patients to participate in it. Also, the true potential impact of the APM will be masked by including providers who are unwilling or unable to successfully implement the care delivery changes that the APM is intended to support. Moreover, even if the APM is successful, that does not mean it would be desirable for every provider to implement it; in many cases, it will likely be both desirable and appropriate to create permanent but voluntary APMs.

D. Revising the APM Parameters

The parameters of the APM (i.e., the amounts paid for individual services or bundles of services, the utilization/spending targets, the quality targets, etc.) will have to be updated regularly to reflect changes in the costs of delivering services, new evidence about the causes and appropriate treatments of diseases, new technologies for diagnosing or treating disease, and changes in the prevalence or severity of health conditions. Failure to do so could mean that the APM would no longer adequately enable and encourage the best quality care at the lowest possible cost. Moreover, healthcare providers may be unwilling to participate if they do not believe appropriate adjustments will be made over time. Once the desired reduction in spending or improvement in quality has been achieved, the Target(s) for the APM would need to change to *maintaining* that lower spending level or improved level of quality.

The creation of an Alternative Payment Model can reveal disparities in the amounts that are being spent for care and the outcomes that are being achieved for that spending that were not visible under the current payment system. To address this without discouraging participation by providers, an APM can begin with customized payments amounts and targets for each provider that are based on the past performance levels of that provider, and then transition over time to payment amounts and targets that are common to all providers or all providers with similar characteristics.

Two fundamentally different approaches can be used to update the parameters of an APM:

- An **analytic approach** that uses analyses of data about costs, outcomes, etc. in an effort to determine what the “right” changes in the APM parameters should be for all providers.
- A **competitive approach** that allows individual providers to determine the prices and Targets based on the costs and outcomes they believe they can achieve, with payers or patients choosing providers based on the parameters they set.

There are advantages and disadvantages to both approaches. In healthcare, analytic approaches and population-level competition are used far more often than patient-level competition, whereas in other industries, the reverse is true.

An effective competitive approach can be developed by combining a well-designed Alternative Payment Model with appropriate mechanisms for transparency and patient cost-sharing. This could be done by:

- setting default parameters using an analytic approach;
- allowing individual providers to set different prices and performance targets;
- allowing patients to choose providers based on prices and quality; and
- updating default parameters based on provider-determined prices and quality targets.

CREATING BETTER VALUE-BASED PAYMENT MODELS

Many current Alternative Payment Models have failed to achieve significant savings or improvements in quality because they have not been designed in ways that will correct the problems created by the current fee-for-service payment. A well-designed APM will:

- pay for the high-value services needed to improve patient care;
- align the amount of payment with the cost of delivering good care;
- assure patients that they will receive appropriate, high-quality care that will achieve a good outcome for them (not just other patients); and
- make the cost of healthcare services more predictable and comparable.

Many current APMs have also had poor results because they fail to preserve four important strengths of the fee-for-service payment system. A well-designed APM will also:

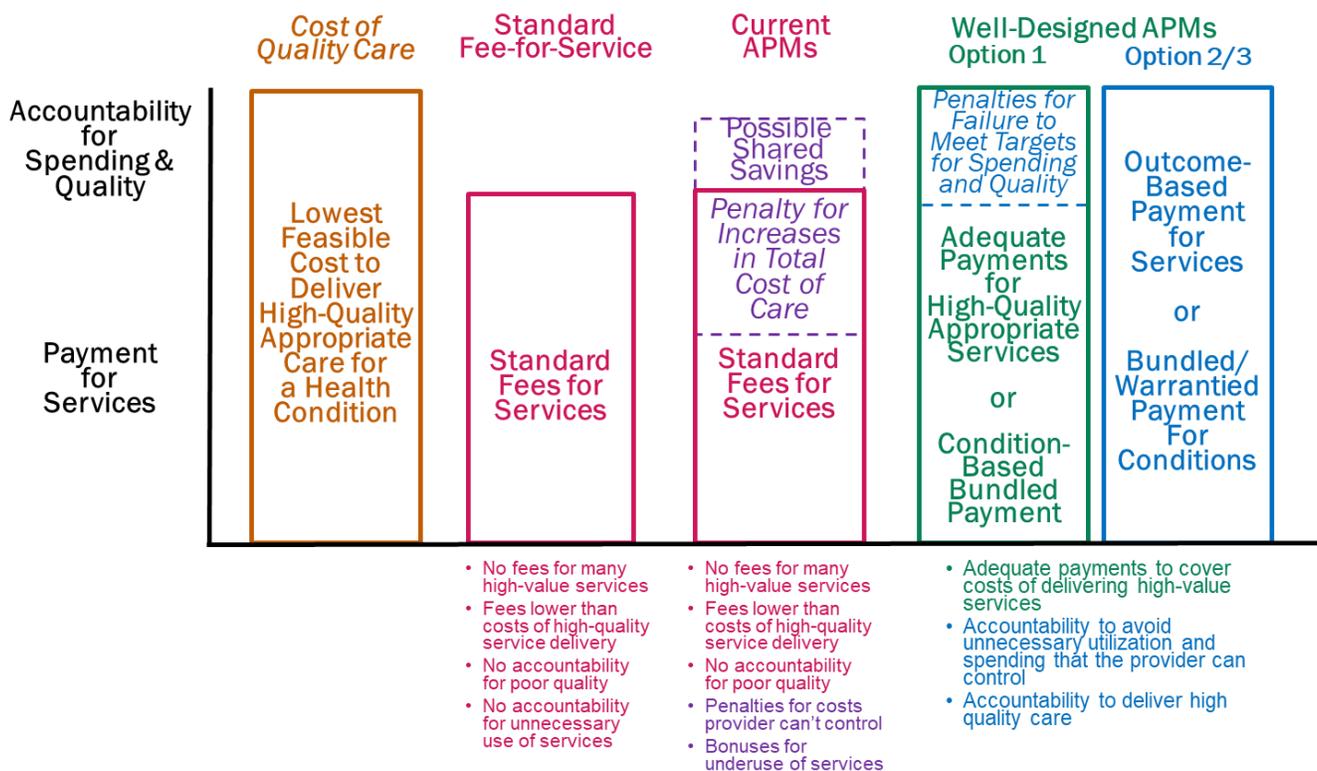
- pay a provider only if a patient receives care;
- make higher payments for patients who need more services;
- base a provider's payment on things the provider can control; and
- enable a provider to know how much they will be paid before delivering a service.

There is no one Alternative Payment Model that will be able to effectively support high-quality care for every type of patient or to effectively address all of the different opportunities for improvement. Multiple, different APMs will be needed. Creating multiple service-specific and condition-specific APMs will not increase fragmentation of care nor will it undercut efforts to improve coordination such as Accountable Care Organizations. In fact, well-designed APMs can help ACOs be more successful than they are today by providing a means to pay the individual providers in the ACO in a way that supports higher-quality, lower-cost care. In contrast, capitation and other "population-based payment systems" simply shift the problems with fee-for-service payments from payers to large provider groups and health systems.

There are also many situations where poor quality of care is caused by underpayment for services and where there are serious risks of losing existing services and seeing outcomes for patients get worse due to inadequate payments. APMs cannot address these problems because spending will need to increase in order to preserve existing services and improve quality and outcomes. Other types of payment reforms will be needed before it is too late to preserve what currently exists.

There is an urgent need to address the high and growing cost of healthcare in America and to do so in a way that improves, rather than worsens, the quality of care for citizens. Alternative Payment Models and other types of payment reforms hold the potential for accelerating progress toward more affordable as well as higher-quality care if, but only if, they are designed in the right way. Faster progress in developing and implementing truly effective healthcare payment systems needs to be a national priority.

STRENGTHS OF WELL-DESIGNED APMs vs. FEE-FOR-SERVICE AND CURRENT APMs



COMPARISON OF WELL-DESIGNED APMs TO CURRENT APMs and FFS

	Current APMs		Well-Designed APMs			
	Shared Savings	Population-Based Payment	Accountable Payment for Services	Accountable Bundled Payment	Outcome-Based Payment	Bundled/Warrantied Payment
Component #1: Adequate Payment for Needed Services	No change in FFS	Flexible payment for each patient; higher amounts for some but not all needs	Payments for new high-value service(s) and/or higher payments for existing service(s)	Bundled payment for group of services from a provider team	Payments for new high-value services and/or higher payments for existing services	Bundled payment for group of services from a provider team
Component #2: Accountability for Spending	Penalty for increase in total cost of care	Fixed payment regardless of services needed or delivered	Penalty if spending controllable by provider exceeds target		Penalty if spending controllable by provider exceeds target	
Component #3: Accountability for Quality	None	Penalties for poor performance on population-level quality measures	Penalty if quality controllable by provider falls short of target for individual patient	Penalty if quality controllable by provider falls short of target for individual patient	No payment if quality standards are not met	Compensation for problems caused by failure to deliver high-quality care
Component #4: Patient Eligibility Determination	Attributed based on service utilization	Attributed based on service utilization	Patient selects provider team	Patient selects provider team	Patient selects provider team	Patient selects provider team

ADDRESSES WEAKNESSES IN FEE-FOR-SERVICE PAYMENT?

Flexibility to deliver all needed high-value services?	NO	YES	YES	YES	YES	YES
Aligns payment with cost?	NO	NO	YES	YES	YES	YES
Assures each patient receives high-quality care?	NO	NO	YES	YES	YES	YES
Makes payments predictable and comparable?	NO	YES	YES	YES	YES	YES

PRESERVES STRENGTHS OF FEE-FOR-SERVICE PAYMENT?

No payment unless a patient receives care?	YES	NO	YES	YES	YES	YES
Higher payments for patients who need more services?	YES	NO	YES	YES	YES	YES
Payment based only on things provider can control?	NO	NO	YES	YES	YES	YES
Provider knows payment before delivering services?	NO	YES	YES	YES	YES	YES

I. WHAT IS AN ALTERNATIVE PAYMENT MODEL?

A. The Need for Alternative Ways of Paying for Healthcare Services

There is broad consensus that current fee-for-service payment systems are a major reason why healthcare spending has grown faster than inflation without any corresponding improvement in the quality of care or patient outcomes. There are four distinct problems with current payment systems that impede the ability to deliver high-quality care at an affordable cost¹:

- **There are no payments at all for many services that can enable higher-quality care to be delivered at a lower cost.** For example:
 - ◆ Physicians are generally only paid for face-to-face visits with patients, even though a phone call or email could help the patient avoid the need for far more expensive services, such as an emergency department visit. Physicians also generally aren't paid for proactive telephone outreach to patients to ensure they get services that could prevent serious health problems or identify problems at earlier stages when they can be treated more successfully and at lower cost.
 - ◆ Primary care physicians and specialists aren't paid for the time they spend communicating with each other to coordinate a patient's care, even though this can avoid ordering duplicate tests and prescribing conflicting medications. Similarly, a physician is not paid for time spent serving as the leader of a multi-physician care team, even if coordination among the physicians would result in better outcomes for the patient.
 - ◆ There is generally no payment for providing palliative care for patients in conjunction with treatment, even though this can improve quality of life for patients and reduce the use of expensive treatments.
 - ◆ There is generally no payment for providing non-health care services (such as transportation to help patients visit the physician's office) which could avoid the need for more expensive medical services (such as the patient being taken by ambulance to an emergency department).
- **Payment rates often differ significantly from the actual cost of delivering high-quality, appropriate care.** In many cases, the payments for healthcare services are much higher than it costs the providers to deliver services; this causes spending to be higher than necessary. However, there are also many cases in which payments are *below* the cost providers incur, particularly if they deliver higher-quality services and do so only when the services are truly needed. Because a high proportion of healthcare costs is fixed in the short run, and because fees are based on average costs, providers are financially rewarded when they deliver unnecessary services and they are financially penalized when they deliver high-quality, appropriate care.

- **There is no assurance that the services a patient receives are appropriate, high-quality, or achieve the results that the patient needs.** In other industries, customers expect products and services to have a warranty against defects and a money-back guarantee of performance. Warranties and performance guarantees reward the producers of high-quality products and services, and they encourage those producers to clearly define the benefits their products and services can and cannot be expected to provide. In contrast, physicians, hospitals, and other healthcare providers are generally paid for delivering services regardless of whether the services are delivered in the highest-quality way, regardless of whether the services have positive or negative effects on the patient, and regardless of whether the services were necessary or appropriate for the patient in the first place.
- **It is impossible for patients or payers to predict the total amount they will need to pay for treatment of a health problem and to compare the amounts across providers.** In other industries, customers know the full price of a product before they buy it and they can compare the prices different manufacturers charge for similar products. In healthcare, patients and payers cannot even obtain an estimate of the combined fees for all of the services they will receive in order to be treated for a health problem, much less receive a guaranteed price for an entire package of services.

All four of these problems contribute to higher-than-necessary healthcare spending and lower-than-desirable quality and outcomes, and unless alternative ways of paying for healthcare are developed that solve these problems, it is unlikely that significant progress will be made in improving the quality and affordability of healthcare services.

B. How MACRA Defines an Alternative Payment Model

The term "alternative payment model" is often used loosely to mean any method of paying for healthcare services that is different from the standard payment methodology. However, in the Medicare program, "Alternative Payment Model" (APM) has a specific meaning that was established by Congress in the Medicare Access and CHIP Reauthorization Act of 2015 (MACRA).²

TABLE 1 REQUIREMENTS FOR ALTERNATIVE PAYMENT MODELS IN MEDICARE

An Alternative Payment Model must:

- Meet the requirements of one of the following provisions of the Social Security Act:
 - ◆ A model under Section 1115A (Center for Medicare and Medicaid Innovation) that
 - improves quality without increasing spending; or
 - reduces spending without reducing quality; or
 - Improves quality and reduces spending
 - ◆ A demonstration under Section 1899 (Shared Savings Program); payment options are:
 - shared savings payments to an ACO;
 - partial capitation payments to an ACO; or
 - other payment models for an ACO that do not result in higher spending for the ACO
 - ◆ A demonstration under Section 1866C (Health Care Quality Demonstration Program)
 - ◆ A required demonstration, e.g.
 - Section 1866D (National Pilot Program on Payment Bundling)
 - Section 1866E (Independence at Home Demonstration Program)
- Require participants to use Certified EHR Technology
 - ◆ Regulations require that 75% of clinicians must be using CEHRT
- Base payment on quality measures comparable to the Merit-Based Incentive Payment System (MIPS)
 - ◆ Regulations require that:
 - at least one measure must have an evidence-based focus and be reliable and valid; and
 - at least one outcome measure must be used unless none are available or applicable

To be an “Advanced Alternative Payment Model,” the “Alternative Payment Entity” participating in the APM must either:

- Bear financial risk for monetary losses in excess of a nominal amount
 - ◆ Regulations require the Alternative Payment Entity to be at risk for paying CMS up to:
 - 3% of spending for which the Entity is responsible under the APM; or
 - 8% of the average total Part A/B revenues of the entities participating in the APM
- Be designated as a medical home expanded under Section 1115A (none currently exist)

Under MACRA, an Alternative Payment Model must meet three criteria. It must:

1. Either be:
 - a. A model under Section 1115A of the Social Security Act (other than a health care innovation award);
 - b. Part of the shared savings program under Section 1899 of the Social Security Act;
 - c. A demonstration under Section 1866C of the Social Security Act;
 or
 - d. A demonstration required by Federal Law.
2. Require participants to use certified EHR technology³; and
3. Provide for payment of covered professional services based on quality measures comparable to those used in the Merit-Based Incentive Payment System.⁴

Although it is not obvious from reading these criteria, an Alternative Payment Model must, in general, either reduce Medicare spending or at least not cause spending to increase beyond what it would have otherwise been.

This is because each of the statutes listed in the first criterion (Sections 1115A, 1899, and 1866C) require that APMs ultimately be “budget neutral,” as explained in more detail below.

APMs Under Section 1115A

Section 1115A was added to the Social Security Act in 2010 by the Patient Protection and Affordable Care Act.⁵ It established the Center for Medicare and Medicaid Innovation (CMMI) to “test innovative payment and service delivery models to reduce ...[Medicare or Medicaid]... expenditures ... while preserving or enhancing the quality of care...”

CMMI is only permitted to test models where “there is evidence that the model addresses a defined population for which there are deficits in care leading to poor clinical outcomes or potentially avoidable expenditures.” In addition, although Section 1115A permits CMMI to implement payment models that will improve quality without generating savings, CMMI is required to “focus” on models that are expected to reduce the costs of the Medicare and/or Medicaid programs while preserving or enhancing the quality of care. Section

1115A also gives CMMI the authority to waive various other requirements of the law if necessary to implement a payment model.

Section 1115A explicitly permits CMMI to test models even if they are not initially “budget neutral,” i.e., an APM could initially result in an increase in Medicare spending. However, the law requires that after testing has begun, an APM must either be terminated or modified unless it is “expected” to meet one of three criteria:

1. to improve the quality of care without increasing spending;
2. to reduce spending without reducing the quality of care; or
3. to improve the quality of care and reduce spending.

The language of MACRA implies that any payment model that CMMI tests (other than a Health Care Innovation Award) is automatically considered an “alternative payment model” for purposes of MACRA.⁶

Payment Models Under Section 1899

Section 1899 was also added to the Social Security Act in 2010 by the Patient Protection and Affordable Care Act.⁷ It created the Medicare Shared Savings Program, which authorizes several different ways of making payments to “Accountable Care Organizations” (ACOs). Under MACRA, the Medicare Shared Savings Program qualifies as an APM.

The law defines an ACO as a group of healthcare providers that:

- is “willing to become accountable for the quality, cost, and overall care of the Medicare fee-for-service beneficiaries assigned to it;”
- has a formal legal structure allowing it “to receive and distribute payments for shared savings ... to participating providers of services and suppliers;”
- has at least 5,000 Medicare beneficiaries assigned to it; and
- includes “primary care ACO professionals that are sufficient for the number of Medicare fee-for-service beneficiaries assigned to it.”

The law authorizes the use of three different ways of making payments to such ACOs:

1. “shared savings,” in which all providers are paid using standard Medicare fee-for-service payment systems, but the ACO can receive an additional payment if (a) it meets quality performance standards and (b) the estimated average per capita Medicare expenditures for beneficiaries assigned to the ACO is a minimum percentage below a spending benchmark;
2. “partial capitation,” in which the ACO is at financial risk for some, but not all, of the fee-for-service spending on the beneficiaries assigned to the ACO; and
3. “other payment models,” which means “any payment model that ... will improve the quality and efficiency of services,” and that “does not result in spending more for such ACO for such beneficiaries than would

otherwise be expended...if the model were not implemented.”

The default payment model under Section 1899 is the shared savings model; Congress authorized, but did not require, implementation of the partial capitation model or other payment models.⁸ The shared savings model does not require that Medicare spending decrease, it simply does not provide any additional payments to an ACO unless spending does decrease.

Payment Models Under Section 1866C

Section 1866C of the Social Security Act was added by the Medicare Prescription Drug, Improvement, and Modernization Act of 2003.⁹ Titled the “Health Care Quality Demonstration Program,” it was originally intended to last for a period of 5 years, but the time limit was removed in 2010 by the Affordable Care Act.

The Health Care Quality Demonstration Program authorizes the use of “alternative payment systems” for “health care groups.” A health care group can be either a group of physicians, an integrated health care delivery system, or an “organization representing regional coalitions of physician groups or integrated delivery systems.” There is no restriction on the nature of the alternative payment system, nor is there a prohibition on making payments that increase Medicare spending initially. However, the law requires that the aggregate expenditures during the entire demonstration period must be no greater than what would have been expended otherwise.

In addition to changes in payment, the Health Care Quality Demonstration Project authorizes modifications to the benefits available to Medicare beneficiaries under Medicare Parts A and B or to the benefits available through a Medicare Advantage plan. It also authorizes the Secretary of HHS to waive other requirements of the Medicare program.

CMS only implemented 3 demonstration projects under the law.¹⁰ However, the authorization to implement additional projects remains in effect. Consequently, Section 1866C could potentially be used to authorize APMs that do not meet the criteria under Section 1115A or Section 1899.

Payment Models Under Demonstrations Required by Federal Law

From time to time, Congress has mandated demonstrations of specific payment models. Payment models established under these demonstrations would qualify as APMs under MACRA. In some cases, there are time limits on the authorization of payment models under these demonstrations. In general, the laws authorizing these demonstrations have required that they be budget neutral.

For example, the Affordable Care Act mandated a National Pilot Program on Payment Bundling (Section 1866D of the Social Security Act) and an Independence at Home Demonstration Program (Section 1866E of the Social Security Act). Section 1866D limits the National Pilot Program on Payment Bundling to a period of 5

years unless it is determined that the program is expected to reduce Medicare spending.¹¹ The Independence at Home Demonstration is structured as a shared savings model similar to Section 1899, so there is no additional spending unless savings have been achieved, but Section 1866E also requires that participants be terminated if they have not achieved savings.¹²

C. Physician-Focused APMs, Advanced APMs, and MIPS APMs

Congress also created provisions in MACRA designed to encourage physicians to participate in APMs and to encourage the creation of APMs in which physicians could participate. This has resulted in several special categories of APMs – “Physician-Focused APMs,” “Advanced APMs,” “Other Payer APMs,” and “MIPS APMs” – that are described in more detail below. However, there is nothing in the law that requires an APM to fit into one of these categories. CMS has the authority to implement APMs that are not physician-focused and that do not qualify as either Advanced APMs or MIPS APMs if it wishes to do so.¹³

Physician-Focused Payment Models

In response to concerns that it was not possible for many types of physicians to participate in the APMs that had been created by CMMI, MACRA created the Physician-Focused Payment Model Technical Advisory Committee (PTAC) to encourage the development of “physician-focused” alternative payment models, including APMs for specialist physicians.¹⁴ MACRA does not specifically define a physician-focused payment model, but instead requires CMS to do so through regulations. Under the regulations promulgated by CMS, a “physician-focused model” is one (a) in which Medicare is a payer, (b) in which “eligible professionals” (which includes clinicians other than physicians) are participants, (c) in which the eligible professionals play a core role in implementing the APM’s payment methodology, and (d) which targets the quality and costs of services that eligible professionals participating in the Alternative Payment Model provide, order, or can significantly influence. CMS also established ten criteria for physician-focused payment models; the PTAC is required to make comments and recommendations as to whether a particular APM meets the criteria.¹⁵

In the first two years that PTAC was able to accept and review proposals (November 2016-October 2018), physicians, medical societies, and other entities submitted 28 proposals for physician-focused APMs, and the PTAC submitted comments and recommendations to the Secretary of Health and Human Service on 19 of the proposals.¹⁶

“Advanced” APMs

In addition, MACRA stated that physicians who participated at a minimum level in APMs that met additional criteria would receive a bonus equal to 5% of their Medicare fee-for-service payments in 2019-2024, receive a higher annual update than other physicians after 2025,

and be exempt from the requirements of the Merit-Based Incentive Payment System. In order to qualify for these incentives:

1. The physician needs to receive at least 25% of their total Medicare payments in 2017-2018 through an “Alternative Payment Entity,” or have 20% of their Medicare patients paid for through such an entity, with higher percentages required in later years (50% of payments or 35% of patients in 2019);

and

2. The Alternative Payment Entity needs to be participating in an Alternative Payment Model as defined in MACRA and it must either:
 - a. “bear financial risk for monetary losses under the APM in excess of a nominal amount;”
 - or
 - b. be designated as a medical home expanded under Section 1115A.

CMS has promulgated regulations implementing this portion of MACRA; these regulations label APMs that meet the criteria for physicians to receive bonuses as “Advanced Alternative Payment Models.”¹⁷ There has been controversy over how the regulations should define “more than nominal financial risk,” and there have been several revisions since regulations were first proposed in 2016. Under the final regulations promulgated in the fall of 2017, an Alternative Payment Entity other than a primary care medical home is considered to be taking “more than nominal financial risk” if the Entity could potentially owe CMS or forgo an amount of payment from CMS equal to either:

- 8 percent of the estimated average total Medicare Parts A and B revenues of participating APM Entities;¹⁸ or
- 3 percent of the expected expenditures for which an APM Entity is responsible under the APM.

Under the regulations, primary care medical homes that have not been expanded under Section 1115A are permitted to meet lower standards of financial risk than other Alternative Payment Entities.¹⁹

A payment model that does not meet these risk requirements can still qualify as an Alternative Payment Model; it simply would not be classified as an “Advanced” APM. Although physicians participating in such an APM would not be eligible for the bonus payments and higher fee updates authorized by MACRA, they could still benefit in other ways, and patients and Medicare could also benefit from implementation of a “non-Advanced” APM. For example, the Independence at Home Demonstration Program qualifies as an APM (because it is a demonstration required by federal law) but it does not qualify as an Advanced APM or a MIPS APM because it does not meet the criteria established by CMS in its regulations.²⁰

Other Payer Advanced APMs

Beginning in 2019, MACRA permits physicians to count participation in certain types of payment models implemented by payers other than Medicare (i.e., Medicaid

and commercial insurance) toward the thresholds need- ed to qualify for the bonuses and higher updates. These payment models must meet criteria for “Other Payer Advanced APMs” established in regulations by CMS which are similar but not identical to the criteria for Medicare Advanced APMs.²¹

MIPS APMs

In its regulations, CMS has also defined “MIPS APMs,” a category that was not defined by Congress in MACRA. CMS exempts physicians participating in a “MIPS APM” from most of the requirements of MIPS (the Merit- Based Incentive Payment System)²² but they cannot receive the 5% bonus between 2019 and 2024 or the higher updates to their fees after 2025. CMS has de- fined a MIPS APM as an APM in which physicians partic- ipate that “bases payment on cost/utilization and qual- ity measures” but does not meet the threshold for “more than nominal financial risk.”²³

D. The Need for More and Better APMs

Despite the need to address the problems with current payment systems, as of 2018, the majority of health- care providers in the country were not participat- ing in an Alternative Payment Model, and most provid- ers had not even had an opportunity to do so because of the small number and narrow focus of the APMs that had been created.

Moreover, most of the Alternative Payment Models that have been created have a similar and very simplistic structure:

- No changes are made in current fee for service payments;
- The payer (CMS or a private health plan) esti- mates whether its total spending on the patients is lower than it would have otherwise expected;
- The providers receive a “shared savings” or “performance-based” pay- ment if spending is below a target level and they may be required to pay a penalty if it is not; and
- The bonus payment is reduced if quality targets are not met.

Most APMs structured in this way have had disappoint- ing results. The largest of the APMs implemented by CMS – the Medicare Shared Savings Program – has used this approach, but instead of achieving savings, it increased Medicare spending every year from 2013 to 2016 and achieved only a small amount of savings in 2017.²⁴ Another large CMS APM that has used a simi- lar structure – the Bundled Payments for Care Improve- ment (BPCI) program – also had a very limited impact

on spending,²⁵ and CMS decided not to continue the BPCI program in its original form.

CMS and other payers have asserted that these APMs have failed to achieve savings because they do not cre- ate enough “financial risk” for the participating provid- ers. Proposed solutions have included requiring provid- ers to accept “downside risk” (i.e., the healthcare provid- ers would be responsible for refunding payments from CMS or other payers if spending was higher than target levels) and creating “population-based pay- ments” in which healthcare providers would be ex- pected to deliver all of the services a patient needs for a fixed monthly or annual payment.

However, there is no evidence that simply increasing the financial risk for physicians, hospitals, and other providers in these APMs would result in greater savings for Medicare or other payers. For example, in the Medi- care Shared Savings Program, ACOs with downside risk saved less money for Medicare in 2017 than did ACOs with only “upside risk” (and neither group of ACOs saved very much money at all).²⁶ Moreover, transfer- ring financial risk to providers can have undesirable results, including loss of access to services for higher- need patients, higher prices due to consolidation of providers, and lower quality of care. For example, the shared savings/shared risk methodology used in the Medicare Shared Savings Program and other CMS APMs can financially reward a healthcare provider for failing to order or deliver services that patients need and it can financially penalize the provider for things (such as increases in drug prices) they cannot possibly control.²⁷

There is no evidence that simply increasing the financial risk for physicians, hospitals, and other providers in APMs would result in greater savings for Medicare or other payers. Fortunately, there are different and better ways to design Alternative Payment Models that can directly address the problems in the fee-for-service system and achieve savings for Medicare and other payers without placing healthcare providers at significant financial risk or causing patients to worry about whether needed care is being withheld for financial reasons.

A more plausible explanation for the failure of current APMs is that the APMs have not actually solved the prob- lems with fee-for-service pay- ment described earlier.²⁸ In particular, most APMs:

- do not actually change the way physicians, hospitals, and other healthcare provid- ers are paid, so most provid- ers are still unable to deliver many kinds of high-value ser- vices that could improve out- comes and reduce spending.
- do not change the amounts paid for individual services, so they do not cor- rect mismatches between payment amounts and costs. Shared savings bonus- es and penalties are based on changes in spending, not on the actual costs of delivering services.
- do not assure individual patients that they will re- ceive appropriate, high-quality care that achieves good outcomes. Most payment models assess whether quality has changed on average for a group of patients, not whether it has improved or worsened for individual patients.
- do not define the total amount that will be paid for services until long after the services are delivered,

which can make payments even less predictable than under fee-for-service systems.

There is nothing in federal law that requires Alternative Payment Models in Medicare to use a shared savings/shared risk methodology, nor is there any research suggesting that this approach is more effective than others. Indeed, the authorizing legislation for the CMS Innovation Center does not limit it to using any specific approach to payment and it does not even mention the term “shared savings.”²⁹ There is nothing in federal law that prohibits APMs from making changes in the fee-for-service payments made to providers; the statute creating the CMS Innovation Center specifically authorizes it to waive other requirements of federal law that would prevent implementation of an innovative APM.

Fortunately, there are different and better ways to design Alternative Payment Models that can directly address the problems in the fee-for-service system and achieve savings for Medicare and other payers without placing healthcare providers at significant financial risk or causing patients to worry about whether needed care is being withheld for financial reasons. There are also opportunities for improving patient outcomes where payment models other than APMs will be needed. This report describes how to design payment reforms that can support more affordable, higher-quality healthcare services and how to successfully implement them.

II.

OVERVIEW OF HOW AN APM IS CREATED

An Alternative Payment Model is not an end in itself, it is a means to an end, with the goal being better care for patients at a lower cost for payers. Instead of starting with a generic payment model design, such as “shared savings” or “population-based payment,” and forcing healthcare providers to use it whether it makes sense or not, a good APM should be specifically designed in a way that will enable and encourage healthcare providers³⁰ to deliver the highest quality care to patients at the lowest possible cost. This can be done through the following process:

- Step 1:** Identify one or more opportunities for reducing spending and/or improving the quality of care;
- Step 2:** Identify changes in care delivery that will reduce spending or improve quality in those opportunity areas;
- Step 3:** Identify the barriers in the current payment system that prevent or impede implementing the improved approach to care delivery;
- Step 4:** Design the Alternative Payment Model so that it will overcome the barriers in the current payment system and assure the delivery of higher-value care;
- Step 5:** Determine how payers and providers can operationalize the APM as easily and quickly as possible; and
- Step 6:** Implement the APM, assess its performance, and make improvements as needed.

A. Identifying Opportunities for Achieving Savings and Improving Quality

Success in any endeavor is more likely if the goals are clearly identified. Step 1 in defining an Alternative Payment Model is to identify specific opportunities to reduce healthcare spending while maintaining or improving the quality of care for patients. There are many such opportunities, such as eliminating unnecessary tests and procedures, reducing infections and complications from procedures, slowing or preventing the progression of diseases, etc. However, no single or simple payment structure can effectively address all of these opportunities, so it will be important to identify the specific opportunities on which the APM will focus.

Most current APMs have not been designed to focus on specific opportunities for reducing avoidable spending or specific areas where patients are experiencing poor outcomes. Instead, the goal of most APMs has been defined as “reducing total spending” or “reducing the total cost of care.”³¹ While this simple, comprehensive-sounding goal may seem ideal from the perspective of a payer, it can be highly problematic for both the

healthcare providers being paid through the APM and the patients they care for. There are three reasons for this:

- The goal of an APM is not just to reduce healthcare spending, but to reduce spending *while maintaining or improving quality*. There is an infinite number of ways that total spending might be reduced, but only some of them represent better-quality care, while others could be harmful for patients. For example, in the Medicare Shared Savings Program, it is impossible to tell whether an ACO achieved savings by reducing unnecessary or necessary services. An APM that targets specific opportunities to reduce spending by improving the quality of care will be much safer for patients than an APM that rewards providers for *any* reduction in healthcare spending.
- It is impossible to make changes in the way providers are paid for the high-value services required to reduce spending or improve quality without knowing which types of services they will need to deliver, and this depends on understanding what aspects of spending and outcomes are the focus of improvement. It is not surprising that most APMs based on shared savings for total cost of care don’t make any changes in the underlying fee-for-service system, because it isn’t clear what changes in care delivery are needed for success.
- No individual physician, hospital, or other provider delivers all of the services any individual patient receives, and so none of those providers can control all aspects of the total cost of care for their patients. For example, the CMS Oncology Care Model determines whether oncologists are successful or not based on whether they reduce total Medicare spending during the period of time their patients are receiving chemotherapy, including spending on services that have nothing to do with the patient’s cancer and even if spending has increased due to price increases on expensive cancer drugs.³² An APM that places providers at financial risk for total healthcare spending can create multiple, serious problems, including patients failing to receive needed care, bankruptcy for the providers, and consolidation of providers into larger organizations that charge higher prices for all of the services they deliver.³³

It seems clear that Congress did not want APMs to be based primarily on whether they reduced total Medicare spending regardless of how savings was achieved. The statute creating the Center for Medicare and Medicaid Innovation explicitly requires that in order for CMMI to test a payment model, there must be “evidence that the model addresses a defined population for which there are deficits in care leading to poor clinical outcomes or potentially avoidable expenditures.”³⁴

Consequently, every APM should be designed around specific opportunities for improving outcomes and/or reducing potentially avoidable spending. Section III describes the major categories in which opportunities to improve outcomes and/or reduce spending are likely to fall.

B. Identifying Needed Changes in Care Delivery

In theory, it would be desirable to stop paying providers based on the number and types of services they deliver and instead pay them solely based on the outcomes they achieve, such as whether they cure a disease or whether they successfully complete a procedure without any complications. However, as a practical matter, the outcomes that providers can achieve depends on whether the amount they are paid is adequate to cover the costs they will incur, and those costs will depend on the types of services they need to deliver and the flexibility they have about how to deliver those services. It doesn't make sense to create a new payment model unless there is reason to believe that it would enable providers to successfully tackle the specific opportunities for reducing spending and improving quality that have been identified as goals of the APM.

Consequently, Step 2 in defining an Alternative Payment Model is to determine (1) what approaches to care delivery could achieve the goals of the APM and (2) how much those care delivery approaches would cost. This does not mean the payment model should *require* that care be delivered in a specific way, merely that the structure and amounts of payment should be defined in a way that will support at least one approach that can achieve the desired outcomes. If no one knows how a particular aspect of spending could be reduced or a particular type of outcome could be improved, it may be more appropriate to use grant-funded demonstration projects to find successful approaches and only then design an APM to sustain and replicate them.³⁵

Section IV describes several key ways in which healthcare services may need to change in order to reduce spending and improve quality, and how changes in the way services are delivered can affect the costs of delivering those services. It also describes the need to establish a clear business case for an Alternative Payment Model before attempting to design one.

C. Identifying the Barriers in the Current Payment System

Step 3 in designing an Alternative Payment Model is to identify any barriers the current payment system creates that impede or prevent delivering the improved approach to care delivery. If there are no such barriers, there presumably is no need for an Alternative Payment Model. Conversely, if there are aspects of the current payment system that prevent or discourage providers from delivering the kinds of lower-cost, higher-quality care identified in Step 2, then an Alternative Payment Model will not be successful in achieving its goals unless it removes or significantly reduces these barriers.

Failure to identify and rectify problems with current payment systems is a major reason why so many alternative payment models and other "value-based payment systems" have been unsuccessful in reducing spending or improving quality. APMs are often described as creating an "incentive" for healthcare providers to reduce spending or improve quality. However, this language implies that healthcare providers are able to make the necessary changes in care delivery but have merely been *unwilling* to do so, and that a financial reward or penalty is needed to overcome their lack of willingness. The reason that so many oppor-

tunities for improvement exist is because the current fee-for-service system creates significant barriers to delivering higher-value care, such as those described in Section I. Most providers will not need any financial incentive to deliver better care to their patients if the APM removes the barriers to doing so.

Section V describes the major ways in which current payment systems create barriers to delivering lower-cost, higher-quality care. The more innovative the approach to care delivery, the more likely it is that there will be multiple barriers to using this approach under current payment systems, and the APM will need to address all of these in order to be successful.

Most current APMs have not been designed to focus on specific opportunities for reducing avoidable spending or specific areas where patients are experiencing poor outcomes. Instead, the goal of most APMs has been defined as "reducing total spending" or "reducing the total cost of care." While this simple, comprehensive-sounding goal may seem ideal from the perspective of a payer, it can be highly problematic for both the healthcare providers being paid through the APM and the patients they care for.

D. Designing the APM

Once Steps 1-3 have been completed, Step 4 is to define the structure of the Alternative Payment Model. This will generally involve defining four distinct, but inter-related components:

APM Component #1

A mechanism for reducing or eliminating barriers in the current payment system (identified in Step 3) in order to allow implementation of improvements in care delivery (identified in Step 2) that can successfully address specific opportunities for savings and improved quality (identified in Step 1).

APM Component #2

A mechanism for assuring patients and payers that the specific aspects of spending targeted by the APM will decrease (if the goal of the APM is to achieve savings) or will not increase (if the goal of the APM is to improve quality);

APM Component #3

A mechanism for assuring that patients will receive equal or better quality of care and outcomes as they would with the kind of care they receive under the current payment system; and

APM Component #4

A mechanism for determining which patients will be eligible for the services supported by the APM.

There are multiple ways to design each of these components, and the decisions made about one component affect the decisions about the others. Section VI describes the major options that are available for each component, and the Appendix describes examples of how the components could be combined to support delivery of a high-value service, treatment of an acute condition, and management of a chronic condition.

E. Operationalizing the APM

Once a desirable structure for the APM has been defined, Step 5 is to determine how payers and providers can operationalize that structure so it can be used efficiently for individual patients on a day-to-day basis. At least in the near term, most APMs will need to operate in parallel with the existing payment system rather than replacing it because not all payers, providers, and patients will be paying for, delivering, and receiving services under the APM. Consequently, an APM will generally be easiest to operationalize if it can use existing billing systems, claims payment systems, and data collection mechanisms to the maximum extent possible.

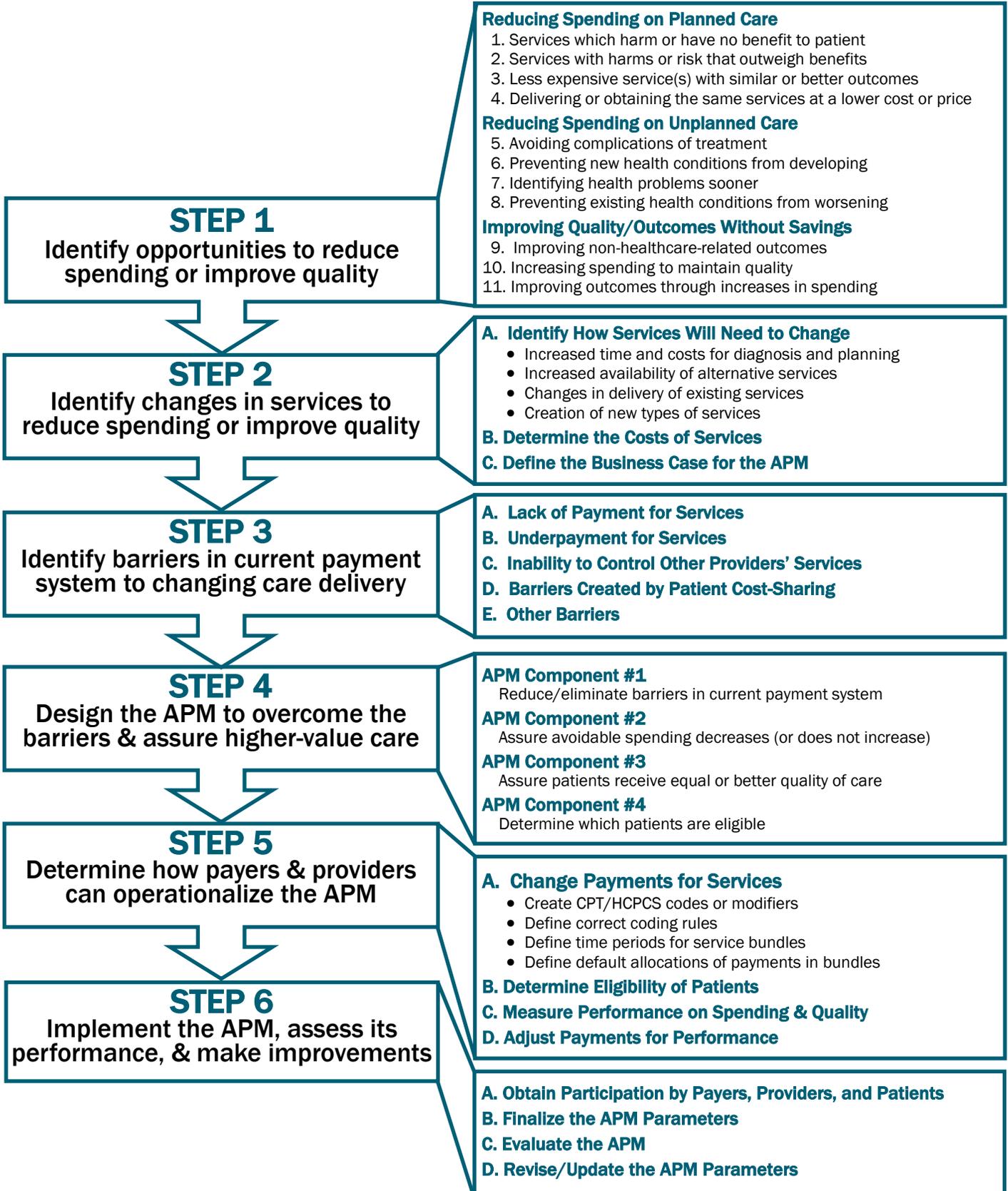
Section VII describes how the structure of an APM can be operationalized within the administrative systems typically used by payers and providers.

F. Implementing the APM

No matter how well an APM is designed, it will not be successful unless it is actually implemented. Moreover, its total impact will depend on how broadly it is ultimately used, and whether it is appropriately adjusted over time to address unanticipated problems and to adapt to changing circumstances.

Section VIII describes the most important actions that will be needed to implement more and better APMs so that as many patients as possible can receive the highest quality care at the most affordable cost.

STEPS TO CREATE AN ALTERNATIVE PAYMENT MODEL



III. IDENTIFYING OPPORTUNITIES FOR SAVINGS & QUALITY IMPROVEMENT

STEP 1

Identify opportunities to reduce spending or improve quality

Reducing Spending on Planned Care

1. Services which harm or have no benefit to patient
2. Services with harms or risk that outweigh benefits
3. Less expensive service(s) with similar or better outcomes
4. Delivering or obtaining the same services at a lower cost or price

Reducing Spending on Unplanned Care

5. Avoiding complications of treatment
6. Preventing new health conditions from developing
7. Identifying health problems sooner
8. Preventing existing health conditions from worsening

Improving Quality/Outcomes Without Savings

9. Improving non-healthcare-related outcomes
10. Increasing spending to maintain quality
11. Improving outcomes through increases in spending

The first step in defining an Alternative Payment Model is to identify specific opportunities to improve outcomes and/or reduce spending that the APM will be designed to address.

Because of the need to control healthcare spending, the priority for Alternative Payment Models will be to achieve reductions in healthcare spending, but to do so in a way that does not harm the quality of care for patients. Opportunities for doing this can be divided into two major categories:

- **Reducing Spending on Planned Care.** One major category of savings opportunities is associated with a conscious decision by a provider of services to change the number and types of services they deliver or order in a way that reduces spending on those services without harming patients. These opportunities represent much of what is often referred to as “waste” in healthcare. This is also the most reliable way that an individual physician or other healthcare provider can generate savings because the provider can plan to make a change in the way they deliver or order services and they can self-monitor to ensure that the change actually occurs.
- **Reducing Spending on Unplanned Care.** There are also many opportunities to achieve savings by reducing the amount of *unplanned care*, i.e., services that are necessary, but only because the patient develops a health condition or a more severe condition that could have been avoided through actions taken at an earlier time. The ability to achieve savings for these opportunities is less certain than with opportunities with changes in planned care because the events causing the need for unplanned care do not occur for all patients. Moreover, some kind of change in *planned care* services will likely be needed in order to achieve the reduction in *unplanned care*; for example, the patient may need to receive a new or different service (e.g., an immunization) that reduces the likelihood of the patient developing a new health condition

(e.g., pneumonia) that would have required expensive treatment. Savings will only occur if the reduction in spending on unplanned care is greater than any increase in spending on planned care needed to achieve it. Determining whether that is the case requires knowing both where the opportunities for savings on unplanned care exist and also what planned services will be needed to achieve those savings.

If the planned or unplanned services that are being reduced are undesirable for the patient, success in reducing spending will frequently be associated with improvements in at least some aspects of the quality of care for the patient.³⁶ However, as will be discussed in Section VI.C, it will be important for the accountability component of the APM to ensure that any changes in planned services do not cause other aspects of quality or outcomes to worsen.

There may also be opportunities to improve the quality of care or outcomes for patients that do not result in any savings. Whether and how these opportunities can be addressed by an APM depends on whether they maintain the current level of spending or increase it:

- **Improving Quality Without Changing Spending.** High spending is not the only problem with the healthcare system, and patients will be better off if quality can be improved even if little or no net savings will result. Under federal law, the Center for Medicare and Medicaid Innovation (CMMI) is explicitly authorized to test models that improve outcomes without any change in spending. (While a change in care delivery may be needed to make the improvements, there may not be a need for an alternative payment model unless the current payment system creates a barrier to making the change.)
- **Improving Quality in Ways That Require Higher Spending.** The fact that spending is too high in general does not mean that it is too high for all patients; some patients may be receiving poor quality care or

achieving poor outcomes because they are not receiving services that would benefit them, but the spending needed to increase services for those patients will not be fully offset by any savings due to the improved outcomes. A payment change designed solely to address these situations will generally not qualify as an “Alternative Payment Model” under federal law because the law requires a payment model to be terminated or modified unless it “is expected to (i) improve the quality of care ...without increasing spending...; (ii) reduce spending...without reducing the quality of care; or (iii) improve the quality of care and reduce spending.” However, this does not necessarily mean that an APM is precluded from pursuing opportunities to improve care that would increase spending; it may be possible, and even desirable in some cases, to combine two changes in services in the same APM – one change that improves outcomes but increases spending, and a second change that reduces spending without harming outcomes. Moreover, if the improvement in quality justifies the increase in spending, the fact that it cannot qualify as an APM simply means that a different approach to payment reform will be needed.

A. Opportunities for Reducing Spending on Planned Care

Opportunities for savings in planned care that maintain or improve the quality of care can be divided into four subcategories:

- Avoiding services that harm or have no benefit for the patient;
- Avoiding services with harms or risk that outweigh the benefits;
- Using a different service or combination of services that is less expensive but achieves similar or better outcomes; and
- Delivering or obtaining the same services at a lower cost or price.

The subcategory in which an opportunity falls is often important for determining how the APM should be structured.

1: Services that harm or have no benefit for the patient

The ideal way to reduce spending without harming quality would be to identify situations in which a service is being used that harms the patient or has no benefit, and then eliminate use of the service in those situations. For example, if a physician orders a treatment that is not effective for the patient’s condition, the spending on that treatment has no benefit for the patient and any side effects of the treatment could create unnecessary harms. If a physician orders a laboratory test or imaging study that will not influence the treatment decision regardless of the result of the test, then avoiding that test or study would reduce spending with no harm to the patient. If the test itself has some harms associated with it – for example, the radiation exposure from an unnecessary CT scan – then avoiding the test both saves money

and improves quality. Ordinarily, these situations will be patient-specific, i.e., the service will benefit some patients but not others. (If the service has no benefits for anyone, there is no reason to deliver or pay for it at all, and it could simply be removed from coverage under insurance rather than requiring creation of an APM.)

Example: Antibiotics have no effect on viral illnesses, yet it has been estimated that 30% of antibiotics prescribed in ambulatory care settings are used for patients who will not benefit from them.³⁷ Not only could spending be reduced without harming the patients by avoiding the use of antibiotics in these situations, in some cases, the patients are actually harmed by taking the unnecessary antibiotics, such as being made more susceptible to developing C. Difficile infections.³⁸ Moreover, reducing unnecessary use of antibiotics would help to slow the development of antibiotic-resistant organisms that require very expensive antibiotics to treat and that can lead to other complications for patients.

Example: Patients who are referred by one physician or hospital to another often have the same tests and imaging studies performed again. This can occur because the second provider was not aware a test was performed previously, because they cannot easily or quickly access the test results, or because fee-for-service payments reward them for carrying out such tests. A study of patients transferred from one hospital to another found that between 42% and 100% of tests performed had been repeated unnecessarily.³⁹

Example: Many patients receive pre-operative testing prior to outpatient surgery even though the results of the tests rarely or never affect the decision to proceed with surgery. One study estimated that Medicare spends \$45 million annually on routine preoperative testing for cataract surgery alone, even though the surgery is very low risk.⁴⁰

Example: Many patients receive imaging for acute low back pain when symptoms first appear, even though no clinical warning signs are present to indicate that such testing is warranted and most cases of back pain will resolve within six weeks following rest or physical therapy. In addition, abnormalities identified during imaging can lead to unnecessary surgery or other undesirable outcomes.⁴¹

There are many other examples with similar characteristics. The Choosing Wisely program⁴² and the U.S. Preventive Services Task Force⁴³ have identified many examples of services that provide little or no clinical benefit and may also cause harm. One study based on an analysis of services in Washington State estimated that 44% of all services received by commercially insured patients were likely unnecessary, representing 33% of all spending.⁴⁴

Misdiagnosis

A subset of situations in which patients receive non-beneficial services deserves special attention. If the patient receives the wrong diagnosis and is treated for that diagnosis, the treatment will at best be unnecessary and at worst harmful. However, in addition, the misdiagnosis may have a variety of other undesirable effects, including higher spending later because the true problem was not treated in a timely fashion, incorrect treatment for other conditions when the treatment decision depends on which other diagnoses the patient has, and errors in risk-adjustment of spending and quality measures.

Example: Two of the most common chronic diseases – asthma and chronic obstructive pulmonary disease (COPD) – are frequently misdiagnosed. In some cases, patients who have one of the diseases are not diagnosed with it, in other cases, patients are diagnosed with one of the diseases when they do not have it, and in still other cases, patients are diagnosed as having COPD when they actually have asthma or vice versa. For example, one study found that 45% of children who had been diagnosed with asthma did not actually have it, while 10% of children with asthma symptoms had asthma but were told they did not.⁴⁵ Another study found that 33% of adults who had been diagnosed with asthma did not actually have the disease, and 2% had a serious cardiopulmonary disease that had been misdiagnosed as asthma.⁴⁶ Yet another study found that 30% of patients who had been told by a physician that they had asthma did not have spirometric testing results consistent with asthma and 87% of patients who had been diagnosed with COPD did not have breathing patterns consistent with COPD when spirometry was performed.⁴⁷ The medications used to treat the diseases are expensive and have undesirable and potentially serious side effects, so treating a patient for one of the diseases when they do not have it is undesirable in terms of both spending and quality.⁴⁸

It is important to recognize, however, that one of the reasons commonly given for overuse of laboratory tests and imaging studies is a desire to avoid misdiagnosis. Consequently, efforts to improve the accuracy of diagnosis could lead to greater overuse of testing and vice versa, so efforts to ensure appropriate testing will likely need to accompany initiatives to improve diagnosis in order to ensure a reduction in net spending.

2: Services with harms or risks that outweigh benefits

In many cases, a service has benefits for patients, but it also has side effects or risks, and the benefits may not outweigh the risks for every patient. Avoiding the use of the service for the subset of patients who would not achieve net benefits can achieve savings without reducing the overall quality of care for patients.

The Choosing Wisely program and clinical guidelines developed by medical specialty societies have identified these kinds of situations as well as the situations dis-

cussed above where a service is simply unnecessary or harmful. It is difficult to determine the magnitude of the opportunities for savings for these kinds of services because they require clinical judgments about a patient and the relevant information needed for that judgment may not be recorded or retrievable even in EHR data much less claims data. However, studies done where clinical data are available have identified examples of expensive services where the risks likely outweigh the benefits for some types of patients who are receiving them.

Example: One of the Choosing Wisely recommendations is to “avoid coronary angiography to assess risk in asymptomatic patients with no evidence of ischemia or other abnormalities on adequate non-invasive testing.”⁴⁹ The only definitive way to determine whether a patient has coronary artery blockage is by performing a cardiac catheterization, but such procedures carry a risk of serious injury or death for the patient as well as a high cost. Consequently, although patients will benefit if a cardiac catheterization identifies an unrecognized coronary artery blockage that can be treated before a heart attack occurs, the risks associated with the procedure will outweigh the benefits for patients who have a very low likelihood of having such a blockage. Making a good decision about whether the risks outweigh the benefits requires a physician to make a careful assessment of the patient using other approaches. A national study of patients that underwent invasive angiography found that the majority (62%) of patients did not have obstructive coronary artery disease and 29% of those patients had no symptoms suggesting that the angiography was warranted.⁵⁰

3: Services that are less expensive but achieve similar or better outcomes

A third way of reducing spending on planned care is to use a different service or combination of services that is less expensive but achieves similar or better outcomes. In some cases, a single service is delivered instead of another service that has a higher cost or price; in other cases, a smaller number of services is delivered. It is possible that the new services are individually more expensive than what would otherwise have been delivered, but if fewer of those services are needed to achieve the same result, the total spent on the full set of services delivered to the patient will be lower than it would otherwise have been.

Example: Many patients will receive better relief from back pain through physical therapy and other non-invasive treatments than through spinal surgery. One study showed that referral of patients to a physiatrist reduced the number of spine operations by 25%, and that patients receiving the physiatry consultation were more satisfied with the results than those who had undergone spinal surgery.⁵¹

Example: Many women with low-risk pregnancies can safely deliver their babies in a birth center rather

than in a hospital. The cost of delivering a baby in a birth center is much lower than in a hospital, and there are also benefits to the mother and baby from doing so.⁵²

Example: Many patients with an acute condition that would ordinarily be treated during an inpatient hospital stay can be treated in their own homes with an intensive home treatment and care program. This “hospital at home” service can cost less than an inpatient hospital stay because it does not require providing room and board to the patients, only treatment services.⁵³

4: Delivering or obtaining the same service at a lower cost or price

Finally, savings can be achieved without any change in the number or type of services if the same services can be delivered by providers who charge less or are paid less for those services. In the Medicare program, where providers of the same type are paid similar amounts to deliver the same service, the largest savings are typically achieved by using a different type of provider, e.g., a service delivered in a physician’s office rather than in a hospital outpatient department. For other payers, however, where prices for the same service can differ dramatically among similar types of providers, savings can also be achieved by referring a patient to a different provider that will deliver the same service at a lower price.

It is important to recognize that savings from the payer’s perspective depends on whether the *price* of the new service is lower than the service that would have been delivered previously; this does not necessarily mean the cost of delivering the new service is lower.

Example: A number of studies have shown that the prices of common outpatient services and procedures are higher when they are delivered in hospitals than when they are delivered in physician offices. One study found that colonoscopies at hospitals cost almost three times as much as in physician offices and chest x-rays cost more than three times as much.⁵⁴

B. Opportunities for Reducing Spending on Unplanned Care

Opportunities for reducing unplanned care can be divided into several subcategories:

- Avoiding complications of treatment;
- Preventing new health conditions from developing;
- Identifying treatable conditions before they worsen; and
- Preventing existing health conditions from worsening.

As with opportunities in planned care, the subcategory in which an opportunity falls is often important for determining how the APM should be structured. In addition, because reductions in unplanned care will likely need to be achieved through changes in some type of planned care, it will generally be important to understand not only the rate at which the unplanned care is occurring overall but the rate at which it is occurring for specific types of patients. Also, in some cases, the unplanned care will occur long after the change in planned services was made, and so the rate at which the unplanned care is being delivered today may overestimate or underestimate the magnitude of the opportunity for improvement in the future.

5: Avoiding complications of treatment

Many patients develop a new health problem because of something that is or is not done while they are being treated for a different health problem. In some cases, commonly called “never events,” there is a direct causal connection between an action or lack of action by a particular clinician and the complication. More commonly, though, the complications only develop in a subset of patients, and there may be ways of reducing the rate of complication but not necessarily of eliminating them entirely.

Example: Hospitalized patients who receive fluids or medications through a catheter inserted into a large vein (a “central line”) are at risk of developing serious infections because of the ability of bacteria to directly enter their bloodstream through the catheter or the insertion site. These central line-associated bloodstream infections (CLABSIs) occur frequently and are very expensive to treat. However, a variety of projects have shown that the rate of such CLABSIs can be reduced dramatically or even virtually eliminated through relatively simple procedures.⁵⁵

Example: Studies have shown that between one-fourth and one-third of Medicare beneficiaries experience adverse events during hospital admissions, skilled nursing facility stays, and rehabilitation hospital stays that result in death, permanent or temporary harm, and prolonged healthcare services. Approximately half of these events were potentially preventable.⁵⁶

6: Preventing new health conditions from developing

“Primary prevention” opportunities involve health conditions that do not result from treatment of other conditions, but rather from the patient’s exposure to disease or environmental conditions or the patient’s lifestyle choices. If the circumstances that cause the condition can be changed, it may be possible to reduce the probability that a patient will develop the condition.

Defining the opportunity means more than merely knowing that a condition can be prevented. It is important to define the probability of successful prevention, how the risk of the condition and the probability of successful prevention varies for different kinds of patients, and how far in the future the preventable event would have occurred.

Example: There is considerable evidence that one of the most common chronic diseases – diabetes – can be prevented in many obese patients if they lose weight.⁵⁷ In 2016, the CMS Office of the Actuary officially certified that the weight losses and lifestyle changes achieved through the Diabetes Prevention Program would result in a reduction in Medicare spending.⁵⁸

7: Identifying health problems sooner

If a condition or complication cannot be reliably prevented in general, or if it was not prevented for a particular patient, it may be less expensive to treat if it is identified at an early stage. This is particularly true for patients who are known to be at risk for specific conditions or complications that will become more serious if there is a delay in treatment; early identification of these problems followed by rapid treatment can both reduce spending and improve patient outcomes.

Example: The U.S. Preventive Services Task Force has recommended routine blood pressure screening for adults in order to identify and treat hypertension before it leads to a stroke, heart attack, or other serious problem.⁵⁹

Example: Many patients with chronic diseases such as asthma, COPD, and heart failure experience exacerbations of their condition that can result in the need for hospitalization if the exacerbation is not identified and treated early enough. One study found that earlier identification and treatment of COPD exacerbations led to faster recovery (approximately one-half day faster recovery for every day sooner the exacerbation was identified and treated) and significantly fewer hospitalizations.⁶⁰

Example: Many patients with cancer are hospitalized or treated in emergency departments for complications resulting from their chemotherapy treatments, such as dehydration or infections. Demonstration projects have shown that many of these ED visits and hospital admissions can be avoided if the symp-

toms of the complications are identified and treated early.⁶¹

It is important to recognize, however, that just because a condition or complication can be treated less expensively at an early stage does not necessarily mean that it is cost-effective to try and identify it at an early stage in order to do so. There are a variety of serious conditions, such as cancer, that can be identified early through screening programs, but if the screening programs have a high false-positive rate (i.e., they inaccurately indicate that a person has the condition when they really do not), the false positive cases will likely result in unnecessary testing or treatments. If the condition for which screening is being performed occurs at a relatively low frequency, the cost of the screening itself and/or the costs of testing/treatment for the false-positive cases may be greater than the savings achieved through early identification and treatment of the “true positive” cases.

Example: Cancer screening programs are not recommended for patients in age groups where the incidence of cancer is low, because with even low false positive rates for the screening test, there may be far more patients inaccurately labeled as having the condition than there are patients who are accurately identified and treated.

8: Preventing existing health conditions from worsening

Finally, even if a condition cannot be cured at an early stage, it may be possible to prevent it from worsening or to slow its progression, thereby avoiding or delaying the more expensive services required to treat more advanced stages of illness.

Example: In most patients, Chronic Kidney Disease (CKD) progresses through several stages until the kidneys can no longer function (i.e., end-stage renal disease, or ESRD) and the patient requires dialysis or a kidney transplant in order to remain alive. However, with proper treatment, the progression of CKD can be slowed, delaying or avoiding the need for expensive ESRD treatments and improving the patients’ quality of life.⁶²

C. Opportunities for Improving Quality or Outcomes Without Savings

If poor quality care results in new or worsened health problems, additional or more expensive healthcare services will likely be needed to address those problems. Consequently, most opportunities to improve the quality of care for patients will also represent opportunities to reduce the utilization of one or more types of healthcare services (either planned or unplanned), so they will fall into one or more of the categories described in the previous two sections. Whether improvements in these aspects of quality produce net savings or not depends on what needs to be done to achieve them and the cost of doing so, which is addressed in Section IV.

However, there are situations in which there are opportunities to maintain or improve outcomes where there will be no savings in healthcare spending or even increases in spending.

9: Improving non-healthcare-related outcomes with no increase in spending

There may be opportunities to improve outcomes for patients that do not directly affect patients' need for other healthcare services and therefore do not affect healthcare spending at all. For example,

- if a patient is unable to work until a health problem is resolved, the sooner the problem can be diagnosed and the sooner treatment is completed, the smaller the negative impact there will be on the patient's income and/or their employer's personnel costs, even if there is no difference in how much is spent on the diagnosis or treatment itself.
- If a particular approach to treating a health problem would be more likely to preserve a patient's ability to engage in recreational activities the patient enjoys, that approach would be preferable in terms of the patient's quality of life even there is no difference in the cost or payment for using that approach.

Improvements in non-healthcare-related outcomes may result in higher incomes or other financial benefits for the patient or others, such as avoiding the need for an employer to hire a temporary worker while an employee is completing treatment, but there might be no savings in healthcare spending. As long as there is no increase in healthcare spending, this approach could qualify as an Alternative Payment Model. However, if healthcare spending had to increase to achieve the improved outcomes, even if the other financial benefits were large enough to offset the increase in healthcare spending, the offsets would not be recorded as savings in healthcare spending and therefore payments to support them would generally not qualify as an Alternative Payment Model.

10: Increases in spending needed to maintain quality

There are situations in which healthcare services that are currently being delivered are achieving desirable outcomes, but the services cannot be sustained at the current amounts of payment. Failure to increase payments could lead to a loss of services and worse outcomes, but increasing payments to maintain the services would result in higher spending.

For example, many physician practices and hospitals in rural areas have closed because they cannot financially sustain their operations at current payment rates. As will be discussed in more detail in Section IV, small and rural providers will generally have higher costs than large, urban providers simply because of the lower volumes of patients, and as populations decrease in rural areas and increase in urban areas, this gap will increase. A payment amount that enables a large provider in an urban area to deliver a service may be inadequate to enable delivery of the same service in a rural area.

Even when special payment programs exist for rural communities, they may not be adequate to cover the costs of services in those areas. For example, through the Critical Access Hospital program, Medicare pays small hospitals in isolated rural areas based on their actual costs of delivering services to Medicare beneficiaries.⁶³ However, although the program originally paid hospitals 101% of their actual costs, under federal sequestration rules, the hospitals are only paid 99% of their costs, no matter how low the costs are. Obviously, if the payment for a service is statutorily required to be lower than the actual cost of delivering the service, the hospital will not be able to sustain the service without finding a source of subsidy, and many small rural hospitals have been unable to do this, particularly since all of their services would require such a subsidy. Increasing payments to Critical Access Hospitals would preserve their ability to provide rapid treatment for patients who are injured or who are experiencing a potential heart attack or stroke, and maternity care for women who are pregnant and ready to deliver, but it would also increase spending on the hospital's services.

In some cases, loss of the rural service could increase spending more than what would be needed to sustain the service, e.g., if delays in treatment due to lack of a local option result in more complications that are more expensive to treat. However, this would at best be an estimate that could never be verified. Moreover, because the rural service is already more expensive than a service in an urban area, an estimate would also have to be made of the savings if patients in the rural areas traveled to urban areas to receive some of the services from providers with lower costs and lower payment amounts.

The fact that these situations do not meet the criteria for an Alternative Payment Model does not mean they cannot or should not be addressed; it simply means that an APM is not the appropriate vehicle for doing so. The implicit assumption behind APMs is that healthcare providers are currently being paid adequately (or more than adequately) for what they are doing today but there are ways of achieving the same or better results with lower spending. If that assumption does not hold – i.e., provid-

ers are not currently being paid enough to sustain high-quality services – then the solution is not to create an APM, but to pay adequately for the services.

11: Improvements in outcomes through increases in spending

Finally, new drugs, medical devices, diagnostic tests, and treatment procedures are constantly being developed that have the potential to improve outcomes for patients but at a higher cost than current tests and treatments. If payment changes are needed to support delivery of these new and improved services, those payment changes would not qualify as an Alternative Payment Model because they would be expected to result in higher spending. As with the previous category, this does not mean that such payment changes should not be pursued, but the changes would need to be made through statutory and regulatory mechanisms other than those governing APMs.

D. Avoiding Increases in Spending

It is important to recognize that “reducing spending” does not necessarily mean that spending will be lower than it is today. The definition of an Alternative Payment Model requires only that spending under the APM be lower than it *would otherwise have been*. Consequently, “savings” can also result from avoiding an increase in spending that would have otherwise occurred. For example:

- If utilization of a service has been increasing and is expected to continue increasing, then slowing or stopping the growth in utilization would generate savings in the future.
- If there are two alternative services that achieve similar outcomes, and access to the less expensive service is decreasing or there is a risk that the alternative service will no longer be offered, actions to preserve access to the less expensive service could avoid an increase in spending in the future.
- If new types of health problems are appearing or if the incidence of existing health conditions is growing, then efforts to reverse those increases could result in lower spending on treatments in the future than would otherwise have been necessary.

These types of savings are often referred to as “bending the cost curve.” Because healthcare spending has historically been increasing so rapidly and consistently, significant savings are possible even if healthcare spending is higher in the future than it is today, as long as the increase is lower than what would otherwise have been expected.

However, quantifying this type of savings is more difficult because it requires making a projection of what utilization and spending on specific services will be in the future. Although total healthcare spending has increased at a consistently high rate for many years, this does not mean that every individual service or aspect of spending will do so. Consequently, there will be greater uncertainty about estimates of savings based on avoiding projected increases in utilization than based on reducing current levels of utilization.

TABLE 2
OPPORTUNITIES TO REDUCE SPENDING AND/OR IMPROVE QUALITY

Opportunity	Examples
Reducing Spending on Planned Care	
1. Avoiding services which harm or have no benefit for the patient	Use of antibiotics for viral illnesses Repeated tests and imaging studies Pre-operative testing prior to outpatient surgery Imaging for acute low back pain Misdiagnosis
2. Avoiding services with harms or risks that outweigh benefits	Coronary angiography in low-risk patients
3. Using services that are less expensive with similar/better outcomes	Physical therapy instead of spinal surgery Low-risk childbirth in a birth center instead of a hospital Home care rather than inpatient admission
4. Delivering/ordering the same service from a provider with a lower cost or price	Diagnostic tests and procedures in physician offices instead of hospitals
Reducing Spending on Unplanned Care	
5. Avoiding complications of treatment	Reducing/eliminating central line-associated bloodstream infections Reducing adverse events during hospital and SNF admissions
6. Preventing new health conditions from developing	Losing weight to prevent diabetes
7. Identifying health problems sooner	Screening for high blood pressure Early identification and treatment of COPD exacerbations Early identification and treatment of chemotherapy complications
8. Preventing health conditions from worsening	Slowing the progression of Chronic Kidney Disease
Improving Quality/Outcomes Without Savings	
9. Improving non-healthcare related outcomes with no increase in spending	Returning patients to work sooner
10. Increasing spending to maintain quality	Increasing payments to sustain primary care physicians Increasing payments to keep small rural hospitals from closing
11. Improving healthcare related outcomes through increases in spending	Expensive new drugs or medical devices that extend life

IV. IDENTIFYING CHANGES IN SERVICES NEEDED TO IMPROVE CARE

STEP 2 Identify changes in services to reduce spending or improve quality

A. Identify How Services Will Need to Change

- Increased time and costs for diagnosis and planning
- Increased availability of alternative services
- Changes in delivery of existing services
- Creation of new types of services

B. Determine the Costs of Services

- Develop a cost model for services
- Identify startup costs
- Consider the cost of time
- Determine whether costs differ for different patients

C. Define the Business Case for the APM

- Will the savings from reductions in avoidable services offset any higher costs of delivering necessary services?

The existence of an opportunity for reducing a particular aspect of spending does not automatically mean that savings in that area can be reliably achieved; there must be a systematic way of delivering care differently that is designed to successfully address that opportunity. Moreover, if the new way of delivering care involves delivering additional or different services, then additional spending on those services must be less than the savings that are achievable in order to produce net savings. To determine whether an APM should be developed and how it should be structured, three separate steps need to be taken:

- a. Identify one or more changes to care delivery that are expected to achieve the desired savings or improvement in quality;
- b. Determine the costs of delivering services under the revised approach to care; and
- c. Determine whether there is a business case for pursuing development of an Alternative Payment Model.

A. Identifying How Services Will Need to Change

After specific opportunities for reducing spending or improving outcomes have been identified as described in Section III, at least one specific set of changes in care delivery should be identified that is expected to successfully achieve improvements in those opportunity areas. Payers often ignore this step; they see an opportunity for reducing spending or improving quality and try to create an “incentive” for healthcare providers to achieve better results without determining whether it is feasible for the providers to do so. By definition, *any* payment model must pay for *something*, and since an *Alternative Payment Model* is expected to achieve a reduction in spending without harming quality or to improve quality without increasing spending, the APM is unlikely to be successful unless it is clear that (a) there is some set of services

that can achieve those results and (b) the APM will pay in a way that enables those services to be delivered.

However, determining that there is at least one feasible way to deliver services that will successfully reduce spending or improve outcomes does not mean that the payment model needs to *mandate* that particular approach. Although it doesn’t make sense to create a new payment model unless there is reason to believe that it would adequately support at least one approach to services that can achieve the goal, it may also be unlikely that a single, specific approach will work equally well in all settings. Section VI will discuss how to design an APM so that it provides the flexibility for different approaches as well as the accountability for achieving the desired savings.

1. Time and Costs of Doing Less

The first two types of opportunities for reducing spending discussed in Section III involve avoiding services that harm patients, have no benefit for the patient, or that create problems or risks that outweigh the benefits. It might appear that these opportunities don’t require delivery of any new or different services, since they are focused on simply avoiding use of an existing service rather than substituting a different service in its place. However, there is a decision-making process involved in determining whether to use a service, and there may well be more time or higher costs associated with deciding *not* to order or deliver a service. For example:

- The physician or other healthcare provider may need to spend more time assessing the patient to determine the correct diagnosis or to determine whether additional testing is needed;
- If the provider needs some type of decision support system to compare the benefits and risks of using a service, there will be time and costs involved in using the system, and there will also be time and costs as-

sociated with contributing data to enable the system to provide more accurate guidance;

- The provider may need to spend time explaining the options to the patient in order to ensure the patient supports the decision not to deliver or order the service; and/or
- The provider may need to spend more time in follow-up contacts with the patient if the service is not delivered or ordered.

As discussed in Section III, avoiding misdiagnosis is an important opportunity that falls into the first subcategory. Studies of the causes of incorrect diagnoses find that in many cases, the physician or other clinician who made the diagnosis needed to spend more time on the patient's case and/or needed better decision support tools.⁶⁴ The short amount of time during standard patient visits makes it hard for a clinician to do a good physical examination and to analyze the reported symptoms in the context of the patient's other characteristics. Moreover, it takes less time to order tests than to determine that a test is not needed, which contributes to overuse of tests and imaging studies.

Example: As mentioned in Section III, there is a high rate of misdiagnosis for patients with asthma and COPD and for patients with symptoms similar to those caused by asthma and COPD. A key reason for misdiagnosis is that many physicians do not use spirometry (or do not perform spirometry correctly) before making a diagnosis; this may be because the physician does not have time to perform the test, because the equipment is not available, or because the physician does not have experience in performing spirometry.⁶⁵

An APM that is designed to improve the accuracy of diagnosis will likely need to support the ability of clinicians to spend additional time on the process of diagnosis. In addition, APMs designed to reduce overuse of testing and imaging will need to ensure that they are not increasing the rate of misdiagnosis.

2. Ensuring Availability of Alternative Services

The third and fourth opportunities discussed in Section III involve replacing a current planned service with an alternative service that is less expensive, more effective, or both. It is not enough to know *what* this alternative service should be; there must also be reason to believe that the service will actually be *available* for the patients who would need to receive it, otherwise the APM will not be successful.

In some cases, the alternative service may not be available at all, and in other cases, the alternative service may be available but there is not currently enough capacity to handle more patients. Almost by definition, the bigger the opportunity to shift care from a service that is cur-

rently being delivered, the less likely it is that there will be adequate capacity to deliver the alternative service,

simply because providers would have had no reason to create capacity that no one would use. If the APM encourages use of the alternative service, then there would be a rationale for providers to create the capacity to deliver it, but it may take time for that capacity to be developed. The APM will need to accommodate that through either or both of the following actions:

- Modifying the accountability targets to reflect the delay in being able to fully achieve them (e.g., requiring a smaller amount of savings initially); and/or
- Increasing the initial payment amounts to reflect the higher costs involved with the initial expansion of services.

As noted in Section III.D, if the alternative service currently exists but there is reason to believe that access to that service will decrease or disappear altogether in the future, actions may be necessary to preserve existing levels of access. Some or all of the savings will come from avoiding a reduction in utilization of the service and a corresponding shift to higher cost services or providers, rather than from increasing utilization of the alternative service compared to current levels.

In some cases, it may not be feasible to deliver the same alternative service to all patients who could benefit from it. For example, in rural areas, the population of the community may not be large enough to make it feasible to offer a particular alternative service, or at least not at the lower cost that made it an attractive target for savings. Many rural communities do not have access to home health services because the long distances between homes make it impractical for home health nurses and aides to manage more than a very small caseload. In these situations, a different approach to care delivery may be needed, and this may require a different APM, or the APM may need to be modified to include different accountability targets that reflect the differences in what is feasible to achieve in different communities.

3. Planned Services Needed to Reduce Unplanned Care

In the second group of savings opportunities discussed in Section III (reductions in spending on unplanned care), it is implicit that one or more changes in planned services will be needed in order to achieve the savings. The specific types of changes in planned services that would be successful will need to be explicitly identified to ensure that the APM provides adequate financial support to implement the changes. As noted earlier, the APM does not need to *require* the delivery of a particular set of services in a specific way, but both payers and providers need to know that there is at least one set of

services that is feasible to deliver and that will achieve the desired goals.

As with opportunities to change planned care, the services needed to reduce unplanned care may not be available or may be different from current services in important ways, so the APM may need to support the creation, expansion, or modification of services in order to achieve the desired results. This includes:

- **Changing the way an existing planned service is delivered.** In some cases, achieving the desired reduction in unplanned services involves changing the way that an existing planned service is delivered. There may be more time or costs involved in delivering the service in a different way, or there may be short term costs and productivity losses while providers learn new approaches and reorganize care delivery processes. The APM will need to ensure the time or costs are adequately supported, similar to what was discussed above with respect to use of fewer services.

Example: Several projects have demonstrated that central line-associated bloodstream infections (CLABSIs) can be significantly reduced or eliminat-

ed through relatively simple techniques, but some of the steps take additional time, and success requires a focused effort to redesign care processes and measure impact.⁶⁶

- **Ensuring availability of services in specific communities or situations.** In other cases, achieving the goal may involve using a different combination of existing services than are used today, but some of those services may not be available in all communities or for all of the patients who could benefit. The APM will need to either support expansion of the services to those communities and patients or establish different goals for savings or quality improvement.

Example: In an urban area, a community oncology practice could provide office-based services (such as IV hydration) to address complications resulting from chemotherapy side effects without the patient having to go to a hospital emergency department. In rural communities, however, the patient's oncologist may be located in a distant city,

**TABLE 3
CARE CHANGES NEEDED TO REDUCE SPENDING AND/OR IMPROVE QUALITY**

Opportunity	Care Changes Needed	Examples
1. Avoiding services which harm or have no benefit for patients 2. Avoiding services with harms or risks that outweigh benefits	Additional time or costs of doing less	Additional time needed during patient visits for shared decision-making about treatment Additional time outside of patient visits to determine which treatment pathway is most appropriate
3. Using services that are less expensive with similar/better outcomes 4. Delivering/ordering the same services from a provider with a lower cost or price	Ensuring availability of alternative services	Creation or expansion of birth centers Expansion of home health and hospice services to rural communities
5. Avoiding complications of treatment 6. Preventing new health conditions from developing 7. Identifying health problems sooner 8. Preventing health conditions from worsening 9. Improving non-healthcare related outcomes	Changes in the way existing services are delivered	Following checklists to prevent infections Remote monitoring of patients for early identification of problems
10. Increasing spending to maintain quality 11. Improving healthcare related outcomes through increases in spending	Ensuring availability of services	Ensuring adequate home health services in rural communities
	Creation of entirely new services	Delivery of "hospital at home" services who need intensive home care for acute illness

and the only option the patient may have for rapid treatment is the hospital emergency department.

- **Creating an entirely new type of service.** If reducing unplanned care requires delivering an entirely new type of service, there will likely not be any current payment for the service at all, so the APM will need to find a way to both start and sustain the service.

Example: The rapid development in wireless technologies holds the potential for innovative approaches to early identification and improved diagnosis of health problems, remote monitoring of health status, in-home treatment and management of conditions, etc. However, current payment systems generally only support in-person, face-to-face interactions between patients and payment amounts are based on the costs of traditional forms of technology, so it may be impossible to offer a service using new technologies even if that service would cost less and/or have better outcomes.

B. Determining the Costs of the Services to Be Delivered

If one or more new or modified services need to be delivered in order to achieve the desired outcomes, it will be essential to estimate the cost of those services. There are two reasons for this:

- If the cost of delivering the new/modified services exceeds the savings that are expected to result, there may well not be a basis for an APM at all.⁶⁷
- If there is an opportunity for net savings, the APM will need to provide payments sufficient to cover the costs of the new or modified services in order for providers to implement the changes.

Providers often ignore this step; they may identify changes in care delivery they would like to make without being clear about how much those changes will cost and whether the changes will reliably achieve savings or improved quality in the opportunity areas identified.

Even if the expectation is that delivering a different combination of *existing* services will achieve savings or better outcomes, it will still be important to determine the cost of delivering those services in the context of the APM. This is because the cost of delivering a service may be very different from what Medicare or other payers currently pay for the service (assuming they pay for it at all). It is entirely possible that:

- the current amount of payment for one or more of the services is *less* than what it will cost to deliver those services to patients in the APM, making it impossible to sustain the services and achieve the savings the services could make possible; or
- the current amount paid for a service is *higher* than what it will cost to deliver the service to the patients in the APM, thereby preventing the maximum amount of savings from being achieved through delivery of the alternative service.

An APM can provide the mechanism for aligning payments with costs more accurately than is possible under the current payment system, but to do so, the costs of the planned services have to be determined.

1. Developing a Cost Model for Services

Cost Accounting vs. Cost-to-Charge Ratios

A healthcare provider will not be able to determine exactly how much it costs them to deliver a current service unless they have a cost accounting system that apportions time and costs to every service they deliver, or unless they carry out a special study to determine what specific costs they incur for the service of interest.

Rather than determining the actual cost of delivering a service, what has typically been done is to estimate the cost using “cost-to-charge” ratios. The amount the provider charges for the service (i.e., the list price, not the amount the provider actually receives from payers or patients) is multiplied by a “cost-to-charge ratio” for that provider to estimate the cost of the service. This is the approach CMS uses in determining the costs of hospital services in order to set payment rates for those services, and it is also the approach used in determining whether an additional “outlier” payment is needed to cover costs that are unusually high and how large that payment should be.

A provider’s cost-to-charge ratio is determined by taking the total costs the provider incurred for delivering a range of services during a particular period of time (e.g., the prior year) and dividing that amount by the sum of the amounts the provider charged patients or payers for all of those services. The denominator is the sum of the provider’s *charges* for each service, i.e., the “full prices” of the services, not the amounts the provider was actually paid by a payer or patient. The cost-to-charge ratio is, in effect, the provider’s average “markup” from cost to price. There may be one overall cost-to-charge ratio (CCR) for the provider, or there may be separate CCRs for different categories of services, e.g., inpatient vs. outpatient services, or radiology vs. laboratory services, if it is feasible to separate costs into those categories.

This approach implicitly assumes that the charge for each service is proportional to its cost. However, if the provider does not know what the actual cost of each service is, it is impossible for them to base the charge for the service on the cost, and so the “cost” estimated by applying an overall cost-to-charge ratio to the charge may or may not have any relationship to the true cost of the service. Many studies have shown that the charge for a service is likely to depend as much or more on what the market will bear than on the cost the provider incurs in delivering it.⁶⁸ Moreover, since providers are rarely paid their full charges for individual services, there is little incentive to insure the charges for individual services are proportional to their relative costs.

Consequently, the cost-to-charge ratio methodology will generally lead to erroneous conclusions about the actual costs of delivering current services, and so information on the actual costs of delivering a current service, either from a cost accounting system or other

source, will be needed to provide accurate information about current services.

Cost Modeling vs. Cost Accounting

However, a cost accounting system or study is also insufficient, because cost accounting can only describe what the costs have been in the past based on the way services were being delivered in the past. If the APM is designed to change the way services are delivered, or if an entirely new service is going to be delivered, a *cost model* is needed, i.e., a method for determining what the future cost of a service will be, including how the cost will differ under different levels of volume, different standards of quality, etc. To do this, a cost model needs to be able to separately identify fixed costs, variable costs, and semi-variable costs.

Separating Fixed Costs, Variable Costs, and Semi-Variable Costs

Although most current payment systems pay the same amount for each service regardless of how often the service is delivered, this does not mean that the cost for a healthcare provider to deliver the service is the same regardless of the number of times the service is delivered. A significant proportion of most healthcare providers' costs is fixed (i.e., these costs will not change even if the number of services provided changes), at least in the short run. This means that the *average* cost of services (i.e., the cost per service or cost per patient) will *increase* when fewer services are provided and the average cost will decrease when more services are provided. This is particularly true of hospitals, which are expected to have emergency rooms, laboratories, surgery suites, and nursing units staffed and ready to go at all hours even if there are no new patients who need them. However, it is also true of physician practices, which still have to cover the same monthly costs of rent, salaries, EHRs, etc. even if fewer patients come to the practice for revenue-producing office visits, procedures, tests, or other services.

Consequently, an estimate of the average costs of services based on *current* volumes of services will be inadequate to determine how costs will change when the volume of services changes significantly.

EXAMPLE: A physician practice is planning to hire a nurse to provide care management services to the practice's patients who have heart failure. The nurse's salary and benefits total \$80,000 per year. If the practice will have 400 patients with heart failure during the coming year, the cost per patient per month of the nurse's time will be \$16.67. However, if the practice only has 300 heart failure patients, the cost per patient per month for the nurse's time at that practice will be \$22.22, or 33% higher. This is because the cost to each practice of employing a nurse is fixed at \$6,667 per month, regardless of the number of heart failure patients the nurse sees.

A relatively small proportion of healthcare costs are truly variable, meaning they change in direct proportion to the number of patients treated or the number of ser-

vices provided. These are costs for items such as drugs, syringes, medical devices, etc. that (1) are only used if there is a patient to treat, (2) represent an out-of-pocket cost to the provider that is using them, and (3) cost the same regardless of the number of times they are used. For example, if a physician practice administers a particular drug to patients and it has to pay an additional amount to purchase the drug each time it is used, then the drug represents a variable cost to the practice.⁶⁹

However, even if the majority of overall costs is fixed, the proportion of fixed and variable costs may differ from patient to patient and treatment to treatment. For example, there will be a much higher variable cost for a hospital to perform knee replacement surgery than to treat a patient for a COPD exacerbation, because the hospital will have to pay an outside vendor for a very expensive prosthetic knee in order to perform the surgery, and the cost of the knee implant will be high relative to the hospital's other costs, whereas the costs of respiratory therapy drugs and supplies for the COPD patient will ordinarily be much less than the cost of a knee implant.

Some costs may be "semi-variable," i.e., the costs will not change when the number of patients or services changes by a small amount, but the costs will change when the number of patients or services changes significantly. Semi-variable costs are challenging for providers and payment systems to deal with because they can cause average costs to increase or decrease significantly when a small change in the volume of services is enough to cross the "break point" where semi-variable costs will change.

EXAMPLE: A hospital unit has 35 patients and is staffed with five nurses in order to maintain a staffing ratio of one nurse for every 7 patients. If the average patient census decreases by 10% (from 35 to 32), the same number of nurses will still be needed to maintain the minimum staffing ratio, so nursing costs will not change, and the cost per patient will increase. However, if the average patient census decreases by 20% (from 35 to 28), the number of nurses could be reduced from 5 to 4 and nursing costs could be reduced by 20%. (Other costs on the unit would still remain fixed – e.g., the unit supervisor, secretary, etc. – so even with the reduction in nurses, the cost per patient will still increase, but by a lower amount.)

Most "fixed" costs are actually semi-variable, but only in the upward direction, meaning that when an increase in the volume of patients or services reaches the capacity of the provider's facilities, equipment, or management team, additional costs will need to be incurred to enable continued growth in volume. However, once the investment is made to create a particular level of capacity, the cost of that investment is fixed until it is paid off, and during that period of time, if fewer services are delivered, the average cost per service will increase.

In the long run, even fixed costs will become variable, e.g., once facilities and equipment have outlived their useful lives, if the volume of services no longer justifies the current capacity, the current facilities and equipment can be replaced with smaller facilities or fewer pieces of

EXAMPLE OF A COST MODEL

Table 4 shows a simple cost model for a hypothetical care management service in a primary care practice. The care managers are nurses that travel to the homes of patients with chronic diseases to help them learn how to manage their health problems. Assume that the care manager is a salaried employee (with total salary and benefits of \$80,000) who can handle a caseload of up to 400 patients. Assume also that if the primary care practice has more than 400 patients who will need services from a care manager, a second care manager will be hired. Assume further that the care manager incurs an average of \$50 in travel expenses for visiting each patient. Finally, assume that the practice incurs \$20,000 in costs each year for office space and secretarial support for the care managers, but these costs will not change unless there are more than 3 care managers.

Figure 1 (which is a graphical representation of the data in the table in Table 4) shows that the cost per patient is very high if there are only a small number of patients in the practice who need the service; initially, the cost per patient decreases rapidly as the number of patients grows, but then the cost per patient becomes more stable. When the number of patients increases beyond a break-point for the semi-variable costs (i.e., there are enough patients to require hiring an additional care manager), the cost per patient increases and then begins decreasing again if the number of patients continues to increase.

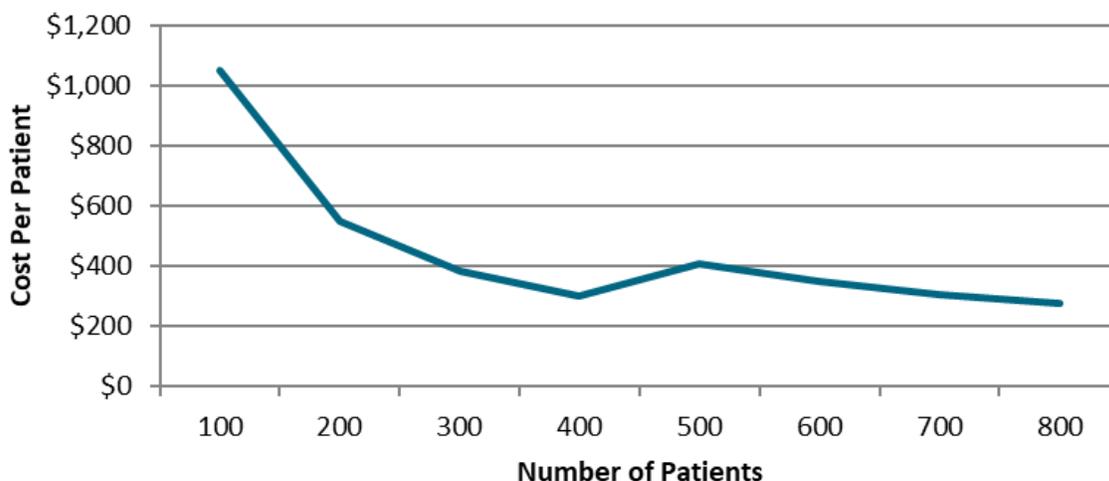
It is important to note that at any point, the marginal cost of delivering the service to additional patients is below the average cost. For example, the data in Table 4 show that with 600 patients, the average cost per patient is \$350, but adding an additional 100 patients to the caseload only adds an additional \$5,000 in cost (the variable cost), or \$50 per patient (not \$350), and the average cost decreases to \$307. Conversely, reducing the number of patients to 500 only reduces costs by \$5,000 (not by \$35,000), and the average cost increases to \$410.

Under a fee-for-service payment system, if the provider were paid \$350 per patient for the service, the provider would break even with 600 patients, make a 14% profit with 700 patients, and have a 15% loss with 500 patients. However, if the service helped patients avoid hospitalizations and resulted in average savings of \$500 per patient, it would be possible to pay more than \$350 and still achieve net savings.

TABLE 4
SERVICE COST PER PATIENT AT DIFFERENT CASELOAD SIZES

Patients:	100	200	300	400	500	600	700	800
Nurse Care Managers	1	1	1	1	2	2	2	2
Fixed Cost (\$20,000)	\$20,000	\$20,000	\$20,000	\$20,000	\$20,000	\$20,000	\$20,000	\$20,000
Semi-Variable Cost (\$80,000, 0-400 Patients)	\$80,000	\$80,000	\$80,000	\$80,000	\$160,000	\$160,000	\$160,000	\$160,000
Variable Cost (\$50/patient)	\$5,000	\$10,000	\$15,000	\$20,000	\$25,000	\$30,000	\$35,000	\$40,000
Total Cost	\$105,000	\$110,000	\$115,000	\$120,000	\$205,000	\$210,000	\$215,000	\$220,000
Cost Per Patient	\$1,050	\$550	\$383	\$300	\$410	\$350	\$307	\$275

FIGURE 1
SERVICE COST PER PATIENT AT DIFFERENT CASELOAD SIZES



equipment, thereby reducing the average cost per service. However, the new facilities and equipment then become fixed costs again. Moreover, the “long run” is generally much longer than the typical length of health insurance policies or contracts between payers and providers, which makes it difficult to ensure that payments will be adequate to cover costs as volume changes.

Modeling Costs Under Multiple Scenarios

Clearly, there is no single number that can describe “the cost” of a service. The cost a provider will incur to deliver a service will *differ* at different levels of patient or service volume, and the average cost per service or per patient will also likely differ depending on how many and what types of patients are receiving services.

2. Identifying Startup Costs

It is rare that any organization can go from one way of delivering services to another way of delivering services without incurring some kind of temporary costs during the transition. For example, if a new employee needs to be hired and trained before a new service can be provided, the provider will incur short-run costs for interviewing, training, and paying initial wages to that employee before the employee can deliver a billable service or achieve the desired benefits for patients. New ways of delivering existing services may temporarily reduce the productivity of existing employees until the new processes are learned and efficiencies are achieved. A new technology may hold significant promise for improving outcomes or productivity, but when it is first used, there will likely be many problems to resolve, and the cost of the technology will likely be much higher than it will be after it is in widespread use.

Unless an APM includes an explicit mechanism to pay for these costs, e.g., through a one-time startup payment, the provider incurring the one-time costs will need to recover them through higher service payments or savings over a period of time. If the new care delivery model and the payment to support it are expected to be in place for several years, then the costs could be amortized over that multi-year period and a small amount could be added to the payments for the services each year to recoup the initial costs over a period of years. However, if payment contracts will only be for one or two years, amortizing upfront costs over that short time period would require a much higher payment for services, which would make it more difficult to demonstrate a positive business case.

3. Considering the Cost of Time

As discussed earlier, in some cases the desired change in care delivery will not involve delivering fundamentally different services, but rather enabling physicians and other healthcare providers to spend more time delivering the “same” service, e.g., a longer office visit with a patient to ensure an accurate diagnosis. The more time that is spent delivering a service, the fewer services that can be delivered during the course of an hour, day, month or other time period. This means the fixed costs of the service provider will have to be recovered from a smaller number of services, the average cost per ser-

vice will increase, and the payment for each service will need to increase correspondingly. Payments for physician services are implicitly or explicitly based on the estimated amount of time the physician needs to spend in delivering the service, so if more time is needed for delivery of a service, the cost of delivering the service will increase.

4. Determining Whether Costs Differ for Different Patients

In addition to determining how costs are influenced by changes in the number of patients treated, it is important to determine whether costs differ for different types of patients. The fact that one patient is receiving the “same” service as another patient does not mean that it will take the same amount of time to deliver the service to each patient or that the exact same supplies or medical devices will be needed in each case.

The only way to ensure that an APM does not underpay or overpay for services when the cost of the services differs significantly for different patients is to determine what patient characteristics affect costs and how the costs differ based on those characteristics. The most relevant characteristics may not be age, sex, or the number and types of diagnoses (which are the characteristics typically used in risk adjustment models) but other clinical and non-clinical characteristics, such as functional status, language proficiency, ability to drive, etc.⁷⁰

5. The Challenges of Estimating Costs Before Services Are First Delivered

Although it is clearly important to understand the costs of delivering services in order to pay for them adequately, it is also difficult to accurately estimate costs before the service is first delivered. In the example above, it may not be clear how many patients a single nurse care manager can manage until some experience has been gained in delivering the services. This creates a “chicken and egg” problem for a new payment model – it will be difficult to accurately determine the right payment amount until the services are delivered, but the services cannot be delivered unless there is an APM to support them. Section VIII.B discusses mechanisms for setting the initial parameters of APMs in order to address this challenge.

In addition, it is possible that new or lower-cost methods of delivering services will appear once it is clear that there is a sustainable mechanism for paying for the services. For example, in services amenable to technology solutions, once the services supported by the APM are being delivered, a firm might develop a new technology that would reduce the cost of delivering the service compared to what is possible with current technologies. In addition, if more providers use a new technology solution, the manufacturer of the technology can sell it for a lower price if it wishes to (since the manufacturer will be able to spread the fixed costs of production across a larger number of products), which would thereby reduce the amount that the provider has to be paid to deliver the service using the technology. Section VIII.D discusses mechanisms for adjusting the parameters of APMs over time in order to address these changes.

C. Defining the Business Case for an Alternative Payment Model

The goal of creating Alternative Payment Models is to reduce healthcare spending while maintaining or improving the quality of care. There is a *business case* for both the payers and providers to design an Alternative Payment Model that will support and encourage implementing the new approach to care delivery if:

1. one or more opportunities to reduce avoidable spending have been identified,
2. at least one method of delivering services has been identified that is expected to successfully reduce that avoidable spending, and
3. the estimated amount of any increase in cost associated with the change in service delivery is less than the savings expected to result from reducing the avoidable spending.

1. Ensuring There is a Business Case for Both Payers and Providers

The business case for the payers depends on saving more on avoidable services/spending than they would spend on new or different services, whereas the business case for the providers depends on the ability to be paid more than the costs of delivering the different mix of services.⁷¹ Assessing the business case from both perspectives requires estimating the expected costs of the services that would be delivered by the participating providers and the savings that would be produced for payers. Business case analyses done from only the payer's perspective or provider's perspective can lead to developing an APM that will likely never be successful:

- Payers often focus solely on the opportunities for savings discussed in Section III. They estimate the amount that could potentially be saved and then propose “incentives” they believe will encourage providers to achieve those savings without ever trying to determine what it would actually cost for the providers to deliver care differently.
- Providers often focus solely on how they would like to deliver care and the higher amounts they would need to be paid for that without determining whether the savings that could be achieved would offset the higher costs and payments.

If both the costs and savings are accurately estimated and there is not a business case for the payer (because the estimated additional cost of the service delivery approach would exceed the expected savings for the payer), there would likely also not be a business case for the provider (because the amount the payer would be willing to pay would be less than the cost of the services the provider would deliver). If there is no business case, then there is no reason to try and design an Alternative Payment Model to support that specific service delivery approach.⁷² In this situation, one or more of the following actions could be taken:

- An effort could be made to find a different approach to service delivery that has similar costs but a larger impact on avoidable spending;

- An effort could be made to find a different approach to service delivery that has a lower cost but a similar impact on avoidable spending;
- An effort could be made to revise the same service delivery approach to reduce its costs.

It is entirely possible that a business case may only exist when the service delivery approach is used by certain types of providers, for certain types of patients, or in certain types of communities. In this case, an Alternative Payment Model may still be desirable, but it would need to be targeted to the specific providers, patients, or communities where the business case does exist. In the example described earlier, if the nursing service is expected to save an average of \$500 per patient in avoided ED visits and hospitalizations, the service would be financially viable in practices with 300 or more eligible patients (because the cost of the service is less than \$500 at those levels of volume), but it would not be viable in smaller practices (because the cost to deliver the service would be higher than the savings it would generate) and it would not be viable if the patients in the practice had early stage chronic diseases with a very low risk of hospitalization (since the savings from any reduction in hospitalizations would likely be very small and less than the cost of the service). The service could well be desirable in terms of the outcomes for the patients, but if total spending increased in order to deliver it, the payments to support it would not qualify as an APM.

As discussed in Section III under Opportunities 10 and 11, there will be some circumstances in which current payment rates are inadequate to sustain existing services and outcomes will worsen unless spending increases, and there will be circumstances in which improvements in outcomes can be achieved, but only by increasing spending. In these situations, there will not be a business case for an Alternative Payment Model, but there may well be a good reason for a different type of payment reform, one in which spending increases and outcomes are improved or prevented for worsening. These other types of payment reform could be implemented simultaneously with APMs that achieve savings in other areas so that overall spending for a payer still decreases. In these cases, the reduction in overall spending might be lower than if the APM alone were implemented, but the improvement in overall outcomes for patients might be greater.

2. Addressing Uncertainty in the Business Case

In many cases, it will likely be difficult to accurately predict either costs or savings or both. If barriers in the payment system have precluded the delivery of desirable services other than in small demonstration projects, there may only be limited experience with how much it costs to deliver new services or to deliver existing services in different ways, and similarly limited experience in the impacts those changes will have on avoidable spending and patient outcomes. In these cases, the business case analysis needs to accurately and objectively reflect this uncertainty without being either overly optimistic or overly pessimistic:

-
- An overly optimistic analysis – i.e., unrealistically low estimates of costs and unrealistically high estimates of savings – could create excessive financial risks for providers and unrealistic expectations for patients and payers, and could divert time and resources away from APMs with a greater chance of success.
 - Conversely, an overly pessimistic analysis – e.g., using worst-case scenarios on both costs and outcomes – could result in a failure to make changes in care delivery that could have significant benefits in terms of lower spending for payers and/or better outcomes for patients.

As will be discussed in more detail in Section VIII.B.2, the only way to reduce the uncertainty about the business case for many APMs will be to “beta test” the APM on a small scale. Although it might seem desirable to avoid this step and focus only on APMs where there is a high degree of certainty that savings will exceed costs, this could well result in the least amount of savings, since the highest certainty about impacts will generally be associated with the most incremental changes in care delivery, and incremental changes in services may only result in small changes in spending or outcomes.

The urgency of the need to slow the growth in healthcare spending and the failure of most current APMs to do so argues for pursuing approaches with a greater potential for savings and quality improvements even if the greater uncertainty about the impacts requires an additional step in the process of testing and implementation. A business case analysis that assesses both the magnitude and causes of uncertainty will help to identify which APMs are worth pursuing.

V. IDENTIFYING THE BARRIERS IN THE CURRENT PAYMENT SYSTEM

STEP 3

Identify barriers in current payment system to changing care delivery

- A. Lack of Payment for Services
- B. Underpayment for Services
- C. Inability to Control Other Providers' Services
- D. Barriers Created by Patient Cost-Sharing
- E. Other Barriers

There are several different ways in which the current payment system can create barriers to implementing the changes in care delivery described in Section IV:

- There may be no payment for one or more of the services that the providers want to deliver;
- The payment for the services to be delivered may be less than needed to cover their costs;
- The providers may be unable to control the types or costs of services delivered by other providers they rely on for a portion of the patients' care; and/or
- The patients may be unable to afford to pay for the services or to pay their share of the cost of services under their insurance plan.

If the APM does not remove these barriers, it will be unlikely to achieve the desired results.

A. Lack of Payment for Services

The current fee-for-service payment system defines specific payment amounts for over 15,000 different services. Despite this, there may be no fee or any other type of payment at all for many of the services that are needed to fully implement the types of care delivery changes described in Section IV. For example:

- Primary care physicians and specialists aren't paid for the time they spend communicating with each other to coordinate a patient's care, even though this can avoid ordering duplicate tests and prescribing conflicting medications.
- Specialists are not paid for phone and email consultations with primary care physicians, even though such communications can result in quicker and more accurate diagnoses and treatment plans, and consultations can avoid unnecessary office visits with the specialist and visits to an emergency department.
- Physicians aren't paid more for spending additional time in a shared decision-making process with patients and family members to explain multiple treatment options, even though this has been shown to reduce the frequency of invasive procedures and low-value treatments.⁷³

- Physicians generally aren't paid to respond to a patient's phone call about a symptom or problem, even though this could help the patient avoid the need for an expensive emergency department visit.
- Physicians generally aren't paid for proactive telephone outreach to patients to ensure they get preventive care services that could prevent serious health problems or identify problems at earlier stages when they can be treated more successfully and at lower cost.
- There may be no payment for services that patients receive from nurses and non-clinician staff, even though providing this type of education and proactive outreach to patients and family members can help patients manage their health problems more effectively and avoid hospitalizations.
- There is no payment for providing palliative care for patients in conjunction with treatment, even though this can improve quality of life for patients and reduce the use of expensive treatments and hospital admissions.
- There may be no payment for providing non-health care services (such as transportation to help patients visit the physician's office) that could avoid the need for more expensive medical services (such as the patient being taken by ambulance to an emergency department).
- Social services for patients are generally not supported through health insurance, even though there is considerable evidence that socioeconomic factors (adequate housing, food, income, etc.) are more important determinants of health outcomes than healthcare services.

Payments for Services Delivered in Some Circumstances But Not Others

In some cases, payments may only be available for a service in certain circumstances, and the eligible circumstances do not include the patients or providers targeted by the APM. For example,

- Medicare pays for care management services for patients with chronic diseases, but only if the patient

has two or more chronic conditions, even though a patient with one serious chronic condition might benefit from care management services more than a patient with two less serious conditions.

- Medicare will pay for palliative care services for a patient with a serious illness only if they meet the eligibility requirements for the Hospice program and are willing to forego treatment for the illness.
- Medicare will only pay for rehabilitation services in a Skilled Nursing Facility if the patient has spent at least three days in a hospital for an inpatient admission.

Why Fill Payment Gaps With an APM Instead of Changing FFS?

It is not surprising that there would be opportunities to reduce spending on the kinds of high-cost services described in Section III if there is no payment for the lower-cost services that would substitute for or avoid the need for them. Healthcare providers can't go bankrupt in order to improve care for patients or save payers money.

Although it might be possible to simply add new payments to the current fee-for-service system to address some of these payment gaps, there are several reasons why creating an APM can be preferable:

- **Mismatch between fee-based revenues and service costs.** As will be described in more detail in the following section, because of the high fixed costs associated with many healthcare services, revenues derived from fixed fees per service will often either be too high or too low relative to the costs of delivering services. Alternative Payment Models provide the opportunity to better align payments with the costs of delivering services.
- **Lack of accountability for utilization or outcomes.** The fact that higher-cost services can be avoided by delivering a particular low-cost service in some circumstances does not necessarily mean that any time the low-cost service is used there will be net savings. Fee-for-service payment systems do not have good mechanisms for ensuring that a particular service will only be delivered when it is necessary or when it most likely to be effective in preventing other problems. In contrast, an Alternative Payment Model can be designed to tie the delivery of a service to the outcomes that it is intended to achieve.
- **Administrative burden and lack of flexibility in service delivery.** In an effort to focus the use of a new service on the types of patients for whom the service is most likely to avoid other kinds of spending, payers may define the service very narrowly, or create prior authorization programs designed to ensure a service is only used in certain circumstances. These approaches can both increase administrative burden on providers (e.g., if they need to document that the service is being used in approved or desirable circumstances) and limit the flexibility that providers have to use the service when it will be effective. For example, relatively few physicians billed Medicare for Chronic Care Management services when the payments were first created because of the restrictions on how services must be delivered and the documentation require-

ments for billing.⁷⁴ In contrast, an Alternative Payment Model can provide greater flexibility with regard to a service by tying payments partially or fully to outcomes.

- **Narrow definitions of budget neutrality.** Under current federal law, adding a new type of service that physicians can bill for under Medicare requires that the payments for all other physician services be reduced so that total spending on all physician services remains the same.⁷⁵ Even if the new service would reduce other types of Medicare spending (e.g., spending on hospital admissions or nursing facility stays.) by more than the spending on the new service, those savings can't be counted toward the budget neutrality calculation if the new service is implemented through a change in the Physician Fee Schedule. Similarly, adding a new outpatient hospital service requires reducing the payments for other outpatient hospital services, even if the new service would reduce spending on inpatient admissions or post-acute care. In contrast, if the new service is paid for under an Alternative Payment Model, then budget neutrality can be calculated considering all types of Medicare services and spending.
- **Inability to test payment changes on a limited scale.** If a new payment for a specific service is added to the fee-for-service payment system, every provider who is qualified to deliver that service would then be able to deliver that service to every patient who could potentially receive it. In contrast, an Alternative Pay-

TABLE 5 Barriers in the Current Payment System

Lack of payment for services

- In all cases
- In some circumstances

Underpayment for services

- In most or all cases
- For new services
- For specific phases of care
- For specific kinds of patients
- Related to volume
 - ◆ Desirable services in rural areas
 - ◆ Reduction in avoidable services
 - ◆ Loss of cross-subsidy for other services

Inability to control/coordinate services delivered by other providers

Barriers created by cost-sharing

- Cost-sharing amounts too high
- Cost-sharing does not reflect differences in value

Other barriers

- Malpractice liability
- Scope of practice laws
- Fraud & abuse laws
- Lack of insurance coverage

ment model creates the ability to pilot test the payment change with a small group of providers and/or patients in order to determine whether the service achieves the desired results, whether there are any unintended consequences, and whether the payment amount matches the costs of delivering the service.

B. Underpayment for Services

1. Underpayment in Most or All Circumstances

In healthcare, there is often little relationship between the price of a service and what it costs to produce it. There are many cases in which the prices of healthcare services are higher than the costs of delivering the services, and one of the goals of an Alternative Payment Model may be to substitute a lower-priced service or provider for one with an unnecessarily high price.

However, there are also cases in which the fees payers pay for services are *below* the cost providers incur in order to deliver a service in a high-quality way to patients who need it. A payment that is too low can be every bit as much of a barrier to delivering a high-value service as no payment at all.

Example: In the Medicare program, a Critical Access Hospital is paid 99% of its costs for the services it delivers to Medicare patients, forcing the hospital to incur a loss on every service, regardless of how efficiently the service is delivered. (Although the statute creating the Critical Access Hospital requires that the hospital be paid 101% of eligible costs attributable to services delivered to Medicare beneficiaries, the 2% reduction in all federal payments under sequestration rules reduced that amount to 99%. In addition, not all costs are considered eligible for payment, and there is no adjustment for the number of uninsured patients the hospital treats, so even without sequestration, a hospital may not have been receiving payments adequate to cover its costs.)

In some cases, underpayment may mean the service won't be delivered at all, but in other cases, the service may be delivered in a lower-quality way.

Example: If Medicare payments to physicians for office visits with patients are not adequate to support the time needed to make an accurate diagnosis, the low payment may not preclude the office visit from occurring, but it may cause spending on lab tests, imaging studies, and specialty referrals to be higher or patient outcomes to be worse.

2. Underpayment for New Services

The underpayment barrier can be particularly large when a provider first starts delivering a new service. As described in Section IV.B, there will often be significant startup costs associated with a new service, or a period of time in which costs have to be incurred before revenue can be generated. A payment amount that is adequate to cover ongoing costs may not be enough to enable recovery of startup costs.

Example: In an independently-owned physician practice, the profit margin on the services and the physician's earnings are one and the same thing, i.e., the physicians do not receive a salary, but they receive whatever is left over after all other practice costs are paid. This means that paying less for a service than it costs to deliver it simply reduces the physician's income. In the early years of the practice, a large part of the physician's income is used to pay off medical school debt, so shortfalls in payment can have a significant negative impact on what the physician has available to spend on housing, food, etc.

3. Underpayment for Specific Phases of Care

The amount of payment may be too low for a service only when it is delivered during certain phases of the care process.

Example: A physician will need to spend more time with a patient when the patient develops a new health problem in order to accurately diagnose the problem and develop an appropriate treatment plan. The Medicare Physician Fee Schedule makes a higher payment for an office visit with a new patient than for an office visit with an "established patient," but for established patients, there is no distinction between the first visit for a new symptom and later visits for a previously treated condition. This means that if a primary care physician refers a patient with a new symptom to a specialist the patient hasn't seen before, the specialist will be paid more to diagnose the symptom than the primary care physician would have been paid.

Example: Most hospitals charge for the nursing care that is provided during a hospital admission using a fixed per diem charge for each day the patient is in the hospital, and many payers still pay a per diem amount for this component of hospital services. However, the intensity of the nursing care during a multi-day hospital stay will ordinarily be much higher during the initial days of the hospital stay than the final days of the stay. Consequently, a fixed per diem payment will likely be lower than the actual cost of care during the initial days of the hospital stay, and higher than the actual cost during the final days of the hospital stay. Under this system, the hospital would be financially penalized for discharging patients quickly.

4. Underpayment for Specific Kinds of Patients

In some cases, the payment is adequate for the service for some types of patients but not for others. If there is only one payment amount for delivery of a service, but the amount of time, staffing, or materials required to deliver the service varies significantly from patient to patient, then the provider will be financially penalized for treating the higher-cost patients.

Example: Under the Medicare Physician Fee Schedule, an oncologist is paid the same amount for an office visit with a new patient with suspected cancer as a family physician is paid for an office visit with a new, otherwise healthy patient with a minor injury or acute condition.

This problem can easily occur where payments are made for “bundles” of services. As will be discussed in more detail in Section VI.A, if most patients need a relatively similar set of services, a single payment for the entire set of services may be administratively simpler, give the provider greater flexibility, and create more predictability about payment and spending than paying separate amounts for each individual service. However, if some patients need more of the services that are included in the bundle, or if the cost of delivering some of the bundled services is higher for those patients, the bundled payment may be less than is needed to support the actual cost of delivering high-quality care to every patient.

Example: Payments for office visits with physicians are intended to cover a group of activities that occur before, during, and after the office visit. This includes an expectation that patients may need a follow-up call after a visit, but patients with complex problems or limited health literacy may need multiple or lengthy calls. The time involved in those calls would go beyond what was assumed in determining the amount of payment for a standard office visit.

5. Underpayment Related to Volume

In some cases, the payment amount for a service will be too low for a subset of providers who deliver the service less frequently than others. Because a significant portion of the costs of many healthcare services is fixed, the average cost of delivering a service will be higher when fewer services are delivered, and so a lower-volume provider of services can experience losses when paid an amount that would be adequate for higher-volume providers.

a. Underpayment for Desirable Services in Rural Areas

This problem can make it difficult or impossible to deliver some kinds of “low-cost” services in rural communities. If the total number of patients who need the service will not generate sufficient revenues at the standard payment rate to cover the cost of delivering the service, then a rural community may not be able to sustain the service even though the standard payment would be adequate to cover costs in urban areas.

Example: Because of the difficulties in attracting and retaining primary care providers in rural areas, Medicare pays Rural Health Clinics (RHCs) based on their actual costs rather than using fixed fees. However, the payments to independent RHCs are capped at maximum amounts that are generally lower than the Clinics’ costs and lower than the amounts paid to physicians in urban areas. The payments for RHCs that are part of Critical Access Hospitals are reduced below the RHC’s actual costs if the primary care physicians fail to achieve a minimum number of patient visits per year, even if that minimum is impossible to achieve in a very small community.⁷⁶

b. Underpayment When Avoidable Services Are Reduced

There may also be an underpayment barrier associated with the higher-cost services that are to be avoided. If savings are achieved by reducing overuse of a particular service, the average cost of delivering that service will increase. Even if the fees paid for the high-cost service are adequate today, they may no longer be adequate when the utilization of that service decreases, and that could jeopardize providers’ ability to deliver the high-cost service to the subset of patients who really need it.⁷⁷

Example: As discussed in Section III, a large proportion of the cardiac catheterizations that are currently performed to diagnose chest pain are probably unnecessary, and eliminating the unnecessary procedures would reduce the total volume of cases significantly. However, cardiac catheterizations are essential for most patients who are experiencing a heart attack, and the hospital’s cardiac catheterization equipment and staff would still need to be ready on a 24/7 basis to deal immediately with heart attacks when they occur. Eliminating the unnecessary procedures would cause the average cost of the necessary procedures to increase significantly, and current payment amounts might not be high enough to sustain the services.

c. Loss of Cross-Subsidy for Other Services

There can be a problem when fewer services are delivered even if the payment remains above the average cost of delivering the services. If the provider has been using the profit margin on one service to offset losses on another service, a lower profit margin on the first service could cause a net loss overall, jeopardizing the provider’s ability to continue offering the other service. For example, many hospitals use the profit margins they generate on orthopedic and cardiac procedures to offset the losses they experience on primary care, maternity care, and mental health services. Providers of all types often use profit margins on services delivered to commercially insured patients to offset losses on Medicaid and uninsured patients.

C. Inability to Control Services Delivered by Other Providers

For all but very simple health problems, most patients will receive two or more separate healthcare services to address the problem. If there is an opportunity to achieve savings by changing the services, a single provider can make the change if all of the services in question are delivered by the same provider (e.g., a single physician or hospital). However, if the services are delivered by different healthcare providers, then it may be more difficult for any one of those providers to make the necessary changes in services. Under current fee-for-service payment systems, only certain types of providers are permitted to deliver specific services, each provider is paid separately for the services they deliver, and an individual provider has only limited control over what other providers deliver and no control over what they are paid. For example:

Example: When a patient receives knee replacement surgery in a hospital, the surgeon is paid for the surgeon's time in planning and performing the surgery, and the hospital is paid separately for the costs of the nurses who assist with the surgery, the knee prosthetic that is used, the supplies that are used during the surgery, the post-surgical nursing care, etc. The amount the hospital spends on knee prosthetics depends on which prosthetics each individual surgeon chooses, so the hospital cannot reduce its charge for the surgery without cooperation from the surgeons. Moreover, if use of a lower-cost prosthetic requires the surgeon to spend more time during surgery, the surgeon won't be able to perform as many surgeries, and the surgeon's fee revenues will decrease even though the change is saving money for the hospital and for payers on each individual surgery.

Example: An orthopedic surgeon may believe that her knee replacement patients would have better outcomes at a lower cost if the skilled nursing facilities and home health agencies that provide post-acute care services to her patients did so in a different way, but the surgeon has no control over whether and how the post-acute care providers are paid and no control over the specific way they deliver services in return for that payment.

Even if two providers find a way to deliver a combined set of services that cost less than what other providers deliver, there is no way for patients or payers to know this or to assure that they will receive the savings. Under the current fee-for-service system, patients and third-party payers typically cannot even obtain an estimate of the combined fees for all of the services the patient will receive for treatment of a particular condition, much less a guaranteed price for the entire package of services.

D. Barriers Created by Cost-Sharing

In some cases, the barrier to delivering higher-value care is not solely the method of payment or the amount a provider is paid for a service, but the amount the patient is required to pay for the service. In health care, a provider's payment generally comes from two separate entities – the patient pays one part (the “cost-sharing amount”) and the rest comes from an insurer or other third-party payer. If the patient feels the cost-sharing amount is unaffordable or is not commensurate with the benefit of the service to them, the patient may not seek out or accept the service even if doing so would enable the insurer to achieve savings on its share of the payments.

A patient's cost-sharing is determined through a complex set of rules. For example:

- if a patient receives a service that is not covered by the health insurance plan, or receives a covered service from an “out-of-network” provider with whom the health plan has no contracted arrangement, the patient may need to pay the full amount the provider charges for the service;
- if a patient has an unmet insurance deductible that is larger than the amount the insurance plan would pay for the service, the patient may have to pay the full amount for the service;
- if a patient has already spent more than a maximum “out of pocket limit” for the year that is specified in their insurance plan, the insurance plan will pay the full amount for a covered service with no cost-sharing required by the patient;

These rules can cause patients to have to pay more or less for the same service at some points in time than others, or to pay a different amount than other patients who are receiving the exact same service. For example, a patient who is treated for a serious illness early in the year may be required to pay less to receive an elective procedure for an unrelated condition than other, healthier patients would have to pay because the first patient would have already met their deductible and/or reached an out-of-pocket spending limit before the cost-sharing amount for the elective procedure is calculated.

The specific barriers caused by cost-sharing depend both on the way cost-sharing rules are defined and also on the nature of the savings opportunity and the way services would be delivered to achieve it.

Impacts of Cost-Sharing Rules on Changes in Planned Care

- If the patient is required to pay the full amount for one service and only a portion of the amount for a higher-cost service, the first service may be more expensive for the patient while being less expensive for the insurance plan. For example, high-deductible health plans can cause lower-priced healthcare services to be as or more expensive for the patient than higher-priced services.
- If the patient needs a non-covered service in order to use or achieve the best outcomes from covered medical services, the lack of coverage may cause overall

spending to increase. For example, if a patient cannot afford transportation to a physician's office, the patient may not be able to receive a preventive service, no matter how adequate the payment is to the physician or how low the patient cost-sharing is for the service itself.

- If the patient is required to pay a fixed copayment for a service or if there is a limit on the total amount of cost-sharing a patient must pay, the patient may have little or no financial incentive to use a lower-cost service or a lower-cost version of the service because the patient would have to pay the same amount for each. Indeed, the patient may actually be encouraged to use the higher-cost version of the service if they believe that higher-priced services have higher quality or better outcomes.
- If the patient's cost-sharing is proportional to the payment amount for the service, then the patient would spend less when a lower-priced service is substituted for a higher-priced service, but the incentive to use the lower-priced service is significantly less than it would be if the patient were paying the full amount.
- It is not always the case that the savings from a change in care occurs immediately, so it is possible that the patient would have to spend more now in order to spend less later. Even if the patient can be assured that the total cost-sharing over a period of time will be less, the patient may not be able to afford to pay a high proportion of the lower amount immediately. For example, if a patient with a health problem has a choice between receiving a one-time medical procedure to treat a problem or taking medications to treat the problem over a long period of time, the cost-sharing amount for the procedure might be lower than the cumulative cost-sharing for the medications, but the patient may not be able to afford the higher short-run cost-sharing for the procedure.

Impacts of Cost-Sharing Rules on Changes in Unplanned Care

Cost-sharing can be a particularly problematic barrier for a service that is expected to reduce the rate at which unplanned care will occur. In general, the probability that an individual patient would have received the unplanned care without the planned preventive service will be greater than 0% but less than 100%. While the patient will definitely have to pay cost-sharing for the high-value service, there is a possibility the patient would have had to pay nothing if they would never have needed the unplanned care the service was designed to avoid. Some patients may choose to gamble, avoiding the high-value service and hoping the unplanned care is not needed, even though that may result in higher overall spending for the payer across a group of patients.

Example: In addition to concerns about the Medicare chronic care management code described earlier, a provider who provides that service has to charge the patient 20% co-insurance for the service each month. The patient may not be willing to pay that extra cost-sharing even though the care management service would reduce Medicare spending by helping the patient stay out of the hospital.

E. Other Barriers

There may also be barriers to delivering the desired services or reducing the avoidable services that have nothing to do with the payment system. For example:

- providers may be unwilling to order fewer diagnostic tests because of a fear of being sued if the test was not performed and the patient is later determined to have the condition the test was designed to detect.
- payers that require prior authorization before certain services can be delivered may refuse to allow a provider to deliver a service to a patient that would have helped avoid the need for other, more expensive services.
- a clinician may be unable to deliver a particular service because the scope of practice laws in the state require that the service only be delivered by individuals with different types of credentials.
- Federal and state fraud and abuse statutes can subject providers to civil and criminal penalties for changing the way staff are compensated or delivering a service to patients for which there is not an explicit payment.
- an individual may not qualify or have access to health insurance coverage for the services they need.

These barriers cannot be addressed by changes in the payment system alone. Although the solutions are beyond the scope of this document, it is important to recognize that an alternative payment model may not be successful, or it may be less successful than hoped, if all of the barriers to better healthcare are not adequately addressed.

VI.

DESIGNING THE ALTERNATIVE PAYMENT MODEL

STEP 4

Design the APM to overcome the barriers & assure higher-value care

APM Component #1

Reduce/eliminate barriers in current payment system

APM Component #2

Assure avoidable spending decreases (or does not increase)

APM Component #3

Assure patients receive equal or better quality of care

APM Component #4

Determine which patients are eligible

Once the steps described in the previous three sections have been completed, the structure of an Alternative Payment Model can be defined. An APM needs four distinct, but interrelated components:

APM Component #1.

A mechanism for reducing or eliminating the barriers in the current payment system that impede delivering the services needed to reduce specific types of avoidable spending.

APM Component #2. A mechanism for assuring patients and payers that the avoidable spending targeted by the APM will decrease (if the goal of the APM is to achieve savings), or that spending will not increase (if the goal of the APM is to improve quality).

APM Component #3.

A mechanism for assuring that patients will receive equal or better quality of care and outcomes as they would with the kind of care they receive under the current payment system.

APM Component #4.

A mechanism for determining which patients will be eligible for the services supported by the APM.

There are multiple ways to implement each of these components. Several options for each are discussed in the sections below.

A. APM Component #1: Removing the Barriers in the Current Payment System

If the current payment system creates barriers to delivering the services needed to achieve reductions in avoidable spending, the APM needs to remove those barriers or at least reduce them. The mechanism used to do that depends on the nature of the barriers and on the ways care may be delivered once the barriers are removed. Fourteen options are described below:

1. Paying for a service that is not currently paid for;
2. Paying for a service through a bundled payment for a group of services;
3. Increasing the payment for a service;
4. Stratifying the payment for a service by the phase of care;
5. Stratifying the payment for a service by patient characteristics;
6. Creating a condition-based payment;
7. Paying to support standby capacity;
8. Making volume-based adjustments to payments;
9. Making additional payments for outlier cases;
10. Paying based on actual costs incurred;

11. Using a multi-component payment structure;
12. Creating a multi-provider bundled payment;
13. Adjusting patient cost-sharing for services; and
14. Creating or changing last-dollar cost sharing for services.

These options are not mutually exclusive, and two or more options may need to be combined, either to address multiple barriers in the current payment system or to avoid creating a new type of barrier by using an overly narrowly-defined payment change. Option 11 describes how multiple payment options could be combined in order to better match the way costs are incurred than any individual option can.

1. Paying for Unpaid Services

If there is no payment at all under the current payment system for a service that is necessary or desirable, there are two basic options that can be used for resolving this problem in an APM:

- Pay a fee when the service is delivered; or
- Include the service in a broader group of services for which a single fee is paid.

Option 1: Pay a Fee When the Service is Delivered

If the barrier to delivering a high-value service is that there is no payment for that service, the most straightforward solution is to simply create a fee for the service. Although people have been led to believe that paying fees for services is inherently bad, it is important to recognize that the fee would only be paid for services as part of the Alternative Payment Model, and so there would also be mechanisms for accountability about spending and quality that would not exist if the new fee were simply added to the standard fee schedule. In particular, the APM would have mechanisms for avoiding overuse of the service that would not exist if a fee for the service were added to the current payment system.

If there are specific circumstances in which delivery of the service is desirable and other circumstances in which it is undesirable, unnecessary, or of low value, then the desirable circumstances can be defined as conditions required in order for the fee to be paid.

There are many current services for which a fee is paid only if the patient has a particular diagnosis or if the service is delivered in a particular circumstance. For example, a physician practice can only receive the Medicare chronic care management payment if the patient has two or more life-limiting illnesses.

If the barrier in the current payment system is that payment for the service is precluded in one or more specific

circumstances, then there is no need to create a new fee for the service, but merely to expand the list of circumstances in which the current fee can be paid. However, if the service would be more or less expensive to deliver in the additional circumstances, then it would be desirable to define a different service with a separate fee because the service would not really be the “same” as what is currently being paid for under other, current circumstances.

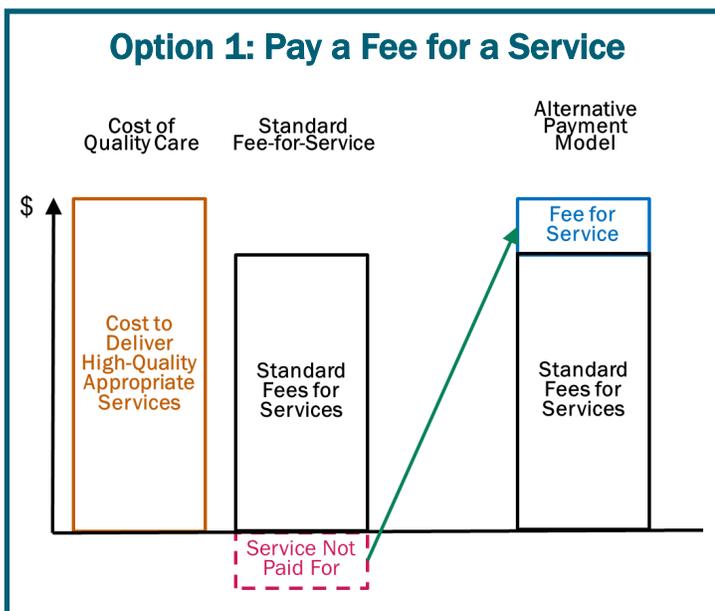
For example, in order to enable more patients to receive healthcare services through telemedicine technologies, Medicare and other payers have begun loosening the rules regarding where patients can be located when they receive telemedicine services and allowing physicians and other providers to bill for additional types of telemedicine services.

In theory, narrowly defining the specific circumstances in which payment will be made can avoid spending money on the service when it will not achieve the expected savings on other services. However, the more complex the definition of the eligible circumstances, the greater the administrative costs that the provider of services will have to incur to document that the criteria for payment have been met, and the greater the administrative costs the payer will have to incur to verify that the documentation is accurate. Even if the payment amount is adequate to cover the costs the provider incurs in delivering the service itself, the payment may not be adequate to also cover the provider’s administrative costs associated with documenting eligibility for payment. For example, several studies have shown that physicians are only billing Medicare for chronic care management services and transitional care services for a small subset of the patients who are likely to be eligible for them due to the documentation requirements involved.⁷⁸ Similarly, even if delivery of the service would achieve savings for the payer in terms of the payments to providers for services, high administrative costs for the payer could reduce or eliminate those savings.

Option 2: Pay for the Service Through a Bundled Fee for a Group of Services

An alternative to paying a separate fee for an individual service is to include the service as a part of a group of services and pay a single “bundled” fee for the group. (A group of services delivered by a single provider is sometimes referred to as a “package” of services, whereas the term “bundle” is often reserved for groups of services delivered by multiple providers.) A bundled fee can be desirable in several different circumstances:

- **if the service should always or almost always be delivered together with the other services in the group.** In this situation, paying for the service separately could result in the service inappropriately being delivered without the others, and it could also result in the service being delivered more or less frequently than necessary, so bundling would help to assure that the appropriate combination of services is delivered.



Example: If a patient with an advanced chronic disease who has been discharged from the hospital (but is not homebound) needs a combination of a home visit by a nurse and an office visit with a primary care physician in order to avoid readmission, but (a) there is no payment for nurse home visits for such patients, (b) nurse home visits alone are less effective if there is no visit with the PCP, and (c) most patients will only need one home visit, then a payment could be made for “a home nursing visit and PCP visit” rather than creating a payment for individual “home nursing visits” and paying separately for the PCP visit. The payment amount would be based on the cost of delivering one PCP visit and an average of slightly more than one home nursing visit (since a small proportion of patients might need two visits).

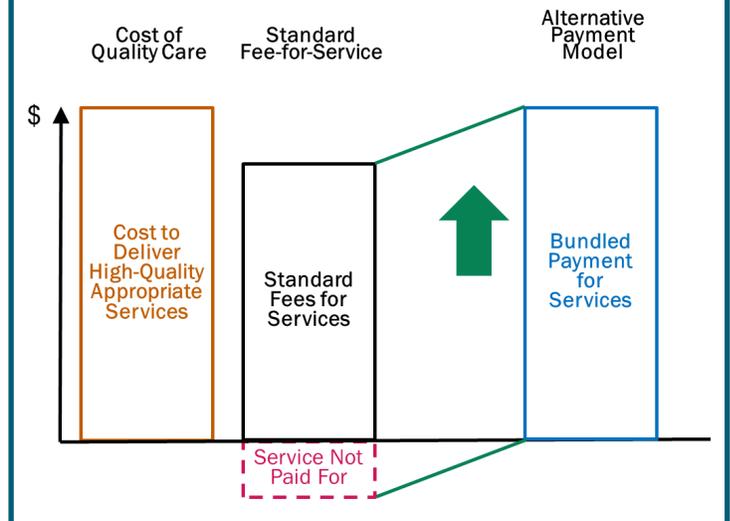
- if the service is intended as an alternative to one or more of the other services in the group. This gives the provider of the services the flexibility to determine which specific services would be delivered, while preventing delivery of both services when one or the other would be sufficient. For example, if either a home visit by a nurse or an office visit with a PCP would reduce hospital readmissions, but the home nurse visits are not currently paid for, then a payment could be created for “nurse home visit or PCP office visit within 30 days after hospital discharge.”
- if there are different ways of delivering the service itself to achieve the same results. If the service has not previously been paid for, then it may be difficult to specify exactly how it should best be delivered. Creating a billing code for a specific service constrains the healthcare provider to deliver the service as described in the code, since billing for the code is a certification that the service associated with the code was delivered. Defining a broader “bundle of services” enables any of the methods of delivering the service to be chosen without encouraging multiple services to be used.

The bundled payment creates more predictable spending for the patient/payer and more predictable revenue for the provider, since the same payment is made regardless of which services or how many services in the bundle are delivered. The provider does not receive less revenue if the patient can be treated with fewer of the services in the bundle or with a lower-cost combination of the bundled services, and the patient and payer do not have to spend more if the provider decides to use more services or a higher-cost combination of services.

Bundled payments are not a new concept. There are several situations in which the current “fee-for-service” payment system pays a bundled payment that is designed to support a particular group of services. Different names are used to describe these payments, including “global fees,” “case rates,” and “service packages,” as well as “bundled payments.” For example:

- Surgeons are typically paid a “global surgery fee” that combines payment for performing the surgery with payment for the visits the surgeon makes with the patient before and after surgery.

Option 2: Create a Bundled Payment



- In the Medicare program and in many commercial insurance contracts, most hospitals are paid for an inpatient admission with a single “case rate” designed to cover all of the services the hospital provides during the patient’s admission. (The amount of the case rate payment is based on the Diagnosis Related Group, or DRG, to which the patient is assigned based on their health conditions and any major procedures they receive in the hospital.)

Structuring the Bundled Payment

Three things have to be specified in order to create a bundled payment:

- **Scope of Services.** Although the bundled payment is intended to provide flexibility as to how many services are delivered, there needs to be a definition of what kinds of services must be delivered in order to qualify for payment and which services will not be paid for separately.
- **Time Period Covered.** Unless all of the services in the bundle are expected to be delivered at exactly the same time, the time period in which delivery of one of the services will be assumed to be part of the bundle must be specified. (For example, global surgery fees typically define a “global period” during which any visits the surgeon makes with the patient are assumed to be included in the bundled fee and cannot be paid separately.) Alternatively, this time period can be defined as the maximum frequency in which the bundled payment can be paid in a particular period of time (e.g., once per week or once per month).
- **Trigger for Payment.** Traditionally, bundled payments have been “triggered” by the delivery of one or more of the services in the bundle. For example, the global surgery fee is paid if and only if a surgery is performed, and the hospital case rate is paid if and only if a patient is admitted to the hospital for an inpatient stay. However, a bundled payment can also be triggered by a health condition, as discussed in Option 6 below.

Bundles Are Not Always Better

Concerns about the problems with traditional fee-for-service payments has led many people to believe that it is always better to pay for services through some kind of a bundled payment. However, bundled payments can be problematic for patients who need more of the individual services contained in the bundle or who need more expensive versions of the services, because even though the provider will incur more costs for delivering more services, the fixed payment for the bundle means the provider will not receive any more revenue to cover the higher costs, and this could result in a patient receiving fewer or less expensive services than they need.

Example: Even though patients receiving chemotherapy for cancer will be receiving both drugs and other services from the oncology practice that is treating them, it is problematic to consider bundling both the drugs and the oncology practice's services together because of the significant variation in costs of the drugs.⁷⁹

Example: Obstetricians typically receive a single bundled payment for maternity care services that covers not only the delivery of the baby, but all prenatal care and post-partum care services for the mother. However, this means that obstetricians will be financially penalized for caring for women with higher-risk pregnancies who need more frequent prenatal care visits.⁸⁰

2. Aligning Payments With the Costs of Services

The second type of barrier described in Section V is when there is currently some kind of payment for a service(s), but the amount of payment is not sufficient to cover the cost of delivering the service(s). Even if the barrier is lack of any payment for a service, creating a new or modified payment as described in the previous section will not be sufficient to enable delivery of a service if the new or modified payment is not adequate to cover the cost of delivering the service.

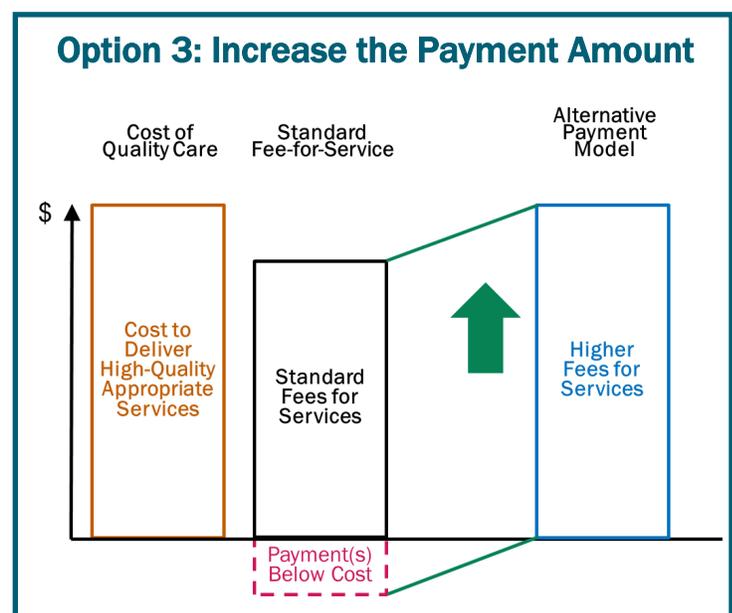
Because there are many different reasons why payments don't match costs, there are many different options for resolving this barrier:

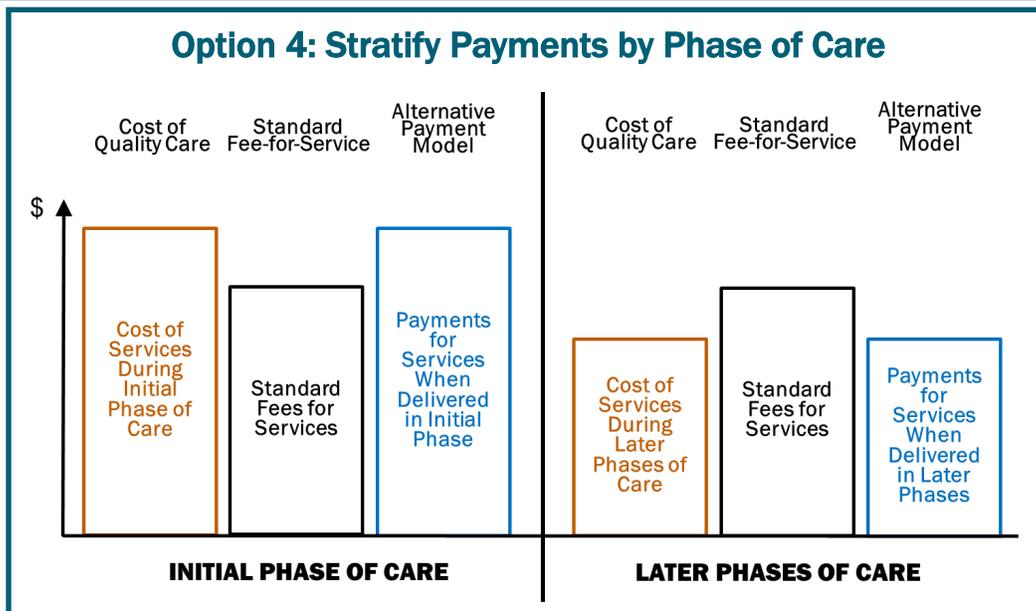
- Increasing the payment amount to cover costs;
- Stratifying payments by phase of care;
- Stratifying payments by patient characteristics;
- Paying based on the patient's condition rather than the services delivered;
- Paying for standby capacity;
- Adjusting payment amounts based on the volume of services;
- Making additional payments for outlier cases;
- Paying based on the costs of care; and
- Using multi-component payment structures

Option 3: Increase the Payment to Cover Costs

If the payment amount for a service is lower than the cost of delivering that service in all or virtually all circumstances, an obvious solution is to increase the amount of payment to match the cost of delivering the service. If the amount of payment for a service has remained unchanged for several years, even though the costs of supplies, utilities, rent, and wages have increased, it may no longer be financially feasible to deliver a service at current payment rates.

However, if there are only specific circumstances in which the current payment is too low, then increasing the payment for the service in all cases would result in an increase in spending for existing utilization and it could result in the service being delivered more frequently than would be desirable because of the higher profit margins that would result. Options 4 and 5 describe two common situations in which payments may be lower than costs in specific circumstances but not in others. In other situations, it may be necessary to define a "different" service (i.e., Option 1) so a higher payment will only be made when the specific circumstances are present that lead to higher costs.





Option 4: Stratify Payments by Phase of Care

As explained in Section V.B, there are situations in which the “same” service or group of services is costlier to deliver in one phase of care than another. To address this, separate payments can be defined for the service in each phase of care, i.e., the amount of payment is determined by both the type of service and the phase of care in which it is delivered. This can be described as “stratifying” the payment by phase of care.

Example: For patients receiving hospice care, Medicare pays a higher per diem amount for Routine Home Care during the first 60 days of care than during subsequent months. During the final 7 days of a hospice patient’s life, Medicare pays a Service Intensity Add-On Payment for services provided by Registered Nurses and social workers in addition to the standard Routine Home Care per diem rate.⁸¹

Example: Standard fee-for-service payments for office visits with physicians are already stratified between the first visit with a patient (a “new patient” visit) and all subsequent visits (“established patient” visits). However, it is reasonable to expect that if a physician and patient are trying to determine the best way to treat a chronic condition, the first several visits after a diagnosis is made will require more time and effort than later visits, because the right medication dosages have to be determined based on the patient’s response, the patient may need education and assistance in using medications or taking other actions to manage their disease, etc. Moreover, if there is uncertainty about the diagnosis, two or more visits may be needed to accurately determine the diagnosis before treatment begins. To address this, three phases of care could be defined with three payment strata – the highest payments for the visits needed to establish a diagnosis and for the initial visits needed to implement and refine the treatment plan, and lower payments for later visits that are used simply to monitor and adjust treatment.

Stratifying the payment does not necessarily mean increasing the payment for some phases while leaving the others the same. The current, un-stratified payment amount for the service may be right “on average” (i.e., when the healthcare provider is delivering services with a typical mix of phases of care), even though the payment amount is too low or too high for each of the individual phases of care. If the payment is increased for one phase of care to better match the higher costs of that phase, it may be possible to reduce the payment amount for one or more other phases. This could mean there would be little or no net increase in spending on the service, or potentially even a reduction in total spending.

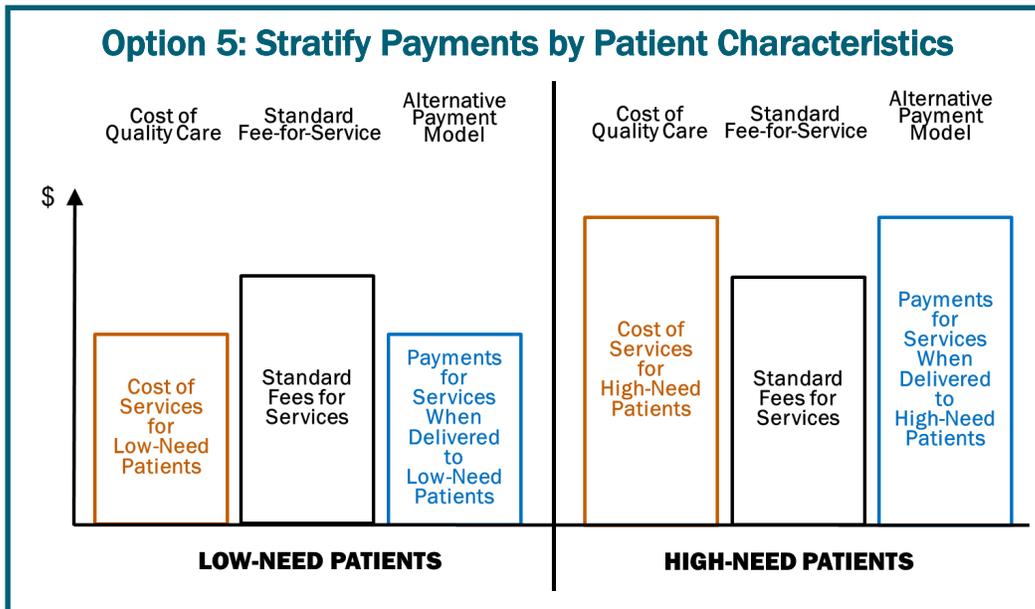
Stratifying payments by phase of care requires defining when a phase of care begins and ends. Many “episode” payments use what are inherently arbitrary lengths of time to define the episode because there is not any clear event that marks an endpoint, particularly in terms of the period of time during which potential complications of a treatment can arise.

Option 5: Stratify Payments by Patient Characteristics

If it takes longer to deliver a service to patients with specific characteristics, or if additional materials or devices are needed for certain types of patients, then the cost of the service will be higher for those patients. In order to cover these higher costs without paying more than is necessary for other patients, higher payments can be defined for the service when it is delivered to patients with specific characteristics.

Most payment systems already implicitly base some payments on patient characteristics by limiting payment to patients with specific types of health problems, but this is just a binary distinction between a payment amount and no payment at all. In contrast, a stratified payment would define two or more different payment amounts for the service based on different patient characteristics.

Option 5: Stratify Payments by Patient Characteristics



Example: It will typically take more time to diagnose or treat a patient if they have difficulty with communication (e.g., speaking a different language, or having cognitive impairments), so payments for a diagnostic or treatment service could be stratified by the patient's communications ability, i.e., a higher amount could be paid for the service when it is delivered to a patient who has communication challenges.

Stratifying the payment could not only involve increasing the payment amount for the patients that require higher-than average costs to serve, but reducing the payment amounts for patients requiring lower-than-average costs. This could mean there would be little or no net increase in spending on the service, or potentially even a reduction in total spending, while also eliminating undesirable incentives for providers to avoid the patients for whom service delivery costs are higher or to deliver the service unnecessarily to the patients for whom the costs of delivering services are much lower.

Stratification vs. Risk-Adjustment

Stratifying payments involves creating two or more discrete categories of patients and assigning a different payment amount to each category. An alternative approach that has been more commonly used in alternative payment models is to “risk-adjust” the payment amount for a service based on characteristics of patients, i.e., to determine a customized payment amount for each individual patient based on their specific characteristics using a regression analysis or other statistical model.

On the surface, risk-adjustment may appear to be more desirable than stratification because

the payment amount can be more finely tuned to small differences in multiple patient characteristics than is possible with discrete categories. However, this seemingly greater precision can actually result in greater inaccuracies in payment for several reasons:⁸²

- The random variation in the cost of delivering services to patients with the same characteristics may be greater than the systematic differences in cost between patients with different characteristics. Most risk adjustment models explain a very small proportion of the variance in spending or outcomes across multiple patients, so for any individual patient, the risk-adjusted payment could be even more “wrong” than payments based on the average costs within a subgroup of patients.
- The statistical models typically used for risk adjustment methodologies assume linear and additive relationships among patient characteristics even though the actual relationships may be non-linear and interactive. For example, a linear model implicitly assumes that the incremental cost or spending for a patient with two problems is equal to the sum of the incremental costs for each of the problems individually, even though it may be either more complicated or less complicated for a provider to treat a patient with multiple conditions depending on what the conditions are.

Instead of using current risk adjustment systems, it will generally be simpler and better to stratify payments based on specific patient characteristics for which there are clinical reasons to expect a causal relationship between the presence of the characteristic and the time the provider will need to spend and/or the number and types of services the provider will need to deliver or order for the patient.

- Most risk adjustment methodologies use variables and weights that are designed to predict the payer's total spending on a patient, not the utilization or spending on a particular service or combination of services to an individual patient. For example, in the Hierarchical Condition Category (HCC) risk adjustment system typically used by CMS in its payment models, the weights

and variables in this system are selected based on their ability to predict total Medicare spending in the subsequent year, and they may or may not have any relationship to the spending on a particular condition. For example, CMS HCC risk adjustment predicts the same spending for a patient with colon cancer as for a patient with diabetes and associated complications, but one would not expect the services provided by an oncologist for the colon cancer patient to be the same as the services provided by an endocrinologist or primary care physician for the diabetic patient.⁸³

- Variables that predict *spending on billable services* may not predict the provider's *actual cost of delivering services*, and the patient characteristics that do affect the provider's cost of delivering a service may not appear to be "statistically significant" in an analysis of the payer's spending on the service. For example, if patients who have more serious conditions have more difficulty coming to the physician's office for appointments, the physician might see them less often but spend more time with them during visits and have more phone contacts with them in between visits; the payer's spending would be lower because there are fewer billable visits, but the provider's cost would be higher because the visits would be longer and there would be more non-billable services delivered. If the current payment amount for a service does not match the actual cost of delivering a service to different patients, an analysis based on spending could lead to variables and weights that are exactly the opposite of what they should be for determining how to risk-adjust payment amounts.
- The fact that patients for whom spending or costs are higher happen to have certain combinations of diagnoses that other patients don't have does not mean that the spending or costs are higher *because* of those diagnoses; the correlation could be purely spurious. However, if higher payments are made for patients with these diagnoses, this creates a perverse incentive to diagnose the patient with those conditions simply to increase the patient's risk score and the associated payment.

Because of these problems, it will generally be simpler and better to stratify payments based on specific patient characteristics for which there are clinical reasons to expect a causal relationship between the presence of the characteristic and the time the provider will need to spend and/or the number and types of services the provider will need to deliver or order for the patient. (Characteristics that affect patient outcomes can also be included in the definitions of strata; this is discussed in Sections VI.B and VI.C.) Moreover, using categorical strata rather than a continuous risk score will mean that payments for two patients will only differ when there is a large enough difference in the characteristics of those patients to justify it. For example, Medicare and most commercial payers stratify payments to hospitals for inpatient services into discrete "Diagnosis Related Groups" rather than paying the hospital a "risk-adjusted" amount that differs for every patient.⁸⁴

Option 6: Condition-Based Payments

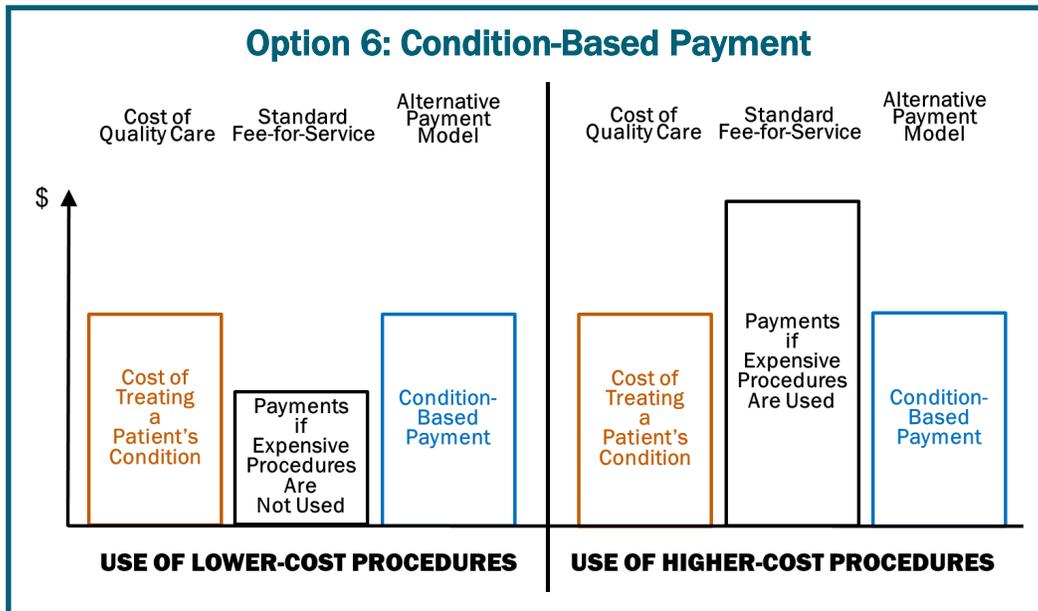
For many services, the cost of delivering the service will depend more on the number and types of patients being treated than the number of times the service is delivered. In these cases, paying based on the number of patients treated for a particular condition – i.e., a "condition-based payment" – would better match the way costs are incurred than paying fees for individual services.

Example: Suppose a primary care practice decides to provide mental health counseling services in order to provide a more integrated approach to patient care:

- **The practice will need to hire one or more mental health counselors to deliver these services. Under the current fee-for-service payment system, the practice will likely be paid a fee for each counseling session that is held, with a higher payment for a longer counseling session. However, the cost to the practice of providing counseling is the salary, benefits, and overhead costs associated with the mental health counselor, and that cost will not change if the counselor delivers one more or one fewer counseling session. If patients fail to attend a scheduled counseling session, which is a common problem in delivering mental health counseling services, the practice would lose money because it is still responsible for paying the counselor even if patients don't come in and fewer payments are received than expected. Paying the practice a monthly amount per patient would better match the way the practice incurs costs than paying based on the number of services delivered.**
- **The number of mental health counselors the practice hires will be based on both the total number of patients receiving primary care from the practice and the proportion of those patients who have behavioral health needs. The practice will need more counselors if it has significantly more patients who have mental health needs, but that will depend on the proportion of patients with mental health needs as well as the total number of patients receiving primary care. As a result, the payment should be condition-based, not just based on the total number of patients the practice sees.**

Condition-based payments will likely not exactly match the way costs are incurred, because costs will generally not change with very small changes in the number of patients (i.e., the costs are semi-variable with respect to patients). However, costs will change when the number of patients increases to the point that a new staff member must be hired, or a new piece of equipment must be purchased. Consequently, a condition-based payment can *better* match the way costs change than paying for each individual service, and it creates different types of incentives for providers. In the example above:

- a per-service payment would result in a reduction in the practice's net revenues if fewer counseling sessions were delivered. In contrast, under a condition-based payment, the revenues would only decrease if the number of patients being managed decreased. The practice would therefore have an incentive to



seek out additional patients who need care in order to maintain adequate revenues, rather than to deliver unnecessary services to current patients.

- if the primary care practice and its mental health counselor can successfully address a patient's behavioral health problem with fewer or shorter counseling sessions, the per-service payment will penalize the practice by reducing its revenues (even though its costs will not change), but the condition-based payment will reward the practice financially if it addresses the patients' needs more quickly and efficiently (because the practice could then accept more patients without having to hire more counselors). The condition-based payment for each patient would be better aligned with a goal of keeping patients healthy than payments based on the number of services delivered.

Using Condition-Based Payments for Diagnosis as Well as Treatment

Condition-based payments can be defined and used in two different ways:

- for **treatment or management of one or more diseases or health problems** after a diagnosis has been established. The "condition" could either be a single disease, a health problem that could lead to disease (e.g., obesity), or a combination of diseases or health problems that can or should be treated, or if they cannot be treated, where the patient requires supportive services; and
- to **determine a diagnosis for patient** who is experiencing a symptom or set of symptoms. In this case, the "condition" is the symptom or combination of symptoms the patient is experiencing.

As discussed in Section III, inaccurate diagnosis represents an important opportunity for avoiding unnecessary services and improving patient outcomes, and as discussed in Section IV, an APM designed to improve the accuracy of diagnosis will need to support the ability of

clinicians to spend adequate time on the diagnostic process. The time and cost involved in making an accurate diagnosis will depend on the nature of the symptoms the patient is experiencing. For some types of symptoms there will be many different potential diagnoses that need to be considered, whereas for other types of symptoms, determining the diagnosis will be relatively straightforward. Moreover, the types of tests needed and the time and costs associated with testing will vary for different symptoms, so different amounts of payment will be needed for different symptoms, and a condition-based payment would allow that.

Defining and Documenting the Condition that Triggers the Payment

Because a condition-based payment is not tied to the number of services delivered the way traditional fee-for-service payments are, there is no financial reward for delivering higher-than-necessary numbers of services. However, since the payment is now triggered by the presence of the condition, there is a risk that patients who do not actually have the condition will be diagnosed as having it so the provider can receive the condition-based payment. Two actions can be taken to minimize this risk:

1. Where it is feasible and cost-effective to do so, the provider can be required to document that the appropriate tests have been performed that document the presence of the condition. This can be done as part of the quality accountability component of the APM (Component #3) and it is discussed in more detail in Section 1.b.v under Component #3.
2. Require that the patient be receiving appropriate treatment or other specialized services for the condition. Since the condition-based payment is intended to cover the costs of treating or managing the condition, it would be inappropriate for a provider to receive the payment if they were not using the payment for services related to the condition. This can be addressed through the utilization accountability

component of the APM (Component #2) and/or the quality accountability component (Component #3).

Stratifying Condition-Based Payments

If patients who have a more severe condition or who have other characteristics will typically need more services, then a provider's per-patient cost for those patients will be higher because fewer patients could be managed or treated by an individual clinician, counselor, nurse, etc. In these cases, a condition-based payment amount that is the same for every patient would penalize the provider for taking on more complex patients whereas a fee-for-service payment system would not. To address this, the condition-based payment could be stratified based on patient characteristics that generally lead to a need for more services, similar to the way service-based payments could be stratified in Option 5. If the per-patient cost of the service is higher during some phases of care than others, then the condition-based payment could be stratified by phase, similar to the approach described for individual services in Option 4.

Bundled Condition-Based Payments

Many patients will need more than one type of service to successfully treat or manage a condition. In addition to using condition-based payment for a single type of service (such as psychotherapy), it could also be used for a bundle of different services (such as psychotherapy and care management). The considerations as to whether to create a bundle of services are similar to those discussed under Option 2. For example, if patients with a particular condition typically need a combination of two different services to treat the condition, the condition-based payment could be designed as a bundle to support delivery of both services. However, if different patients will need very different types or combinations of services, it may be preferable to have separate condition-based payments for different types of services or to stratify the payments for different combinations of the conditions.

Example: Medicare pays most hospitals for medical admissions using a bundled, stratified, condition-based payment – the patient is assigned to a Diagnosis-Related Group (DRG) based on their principal diagnosis and comorbidities, and a single amount is paid to the hospital based on that DRG for all of the services the patient receives during the hospital stay.⁸⁵ In the Medicare Bundled Payments for Care Improvement (BPCI)-Advanced APM, an even larger bundle of services is defined based on some of these same DRGs, with post-acute care services and hospital readmissions included as well as the initial hospital admission.⁸⁶

Defining the Time Period for a Condition-Based Payment

Most individual services take a relatively short amount of time to complete, so the time period covered by a service-based payment is naturally defined by the length of time it requires to complete the delivery of the service. In contrast, a condition-based payment is not tied

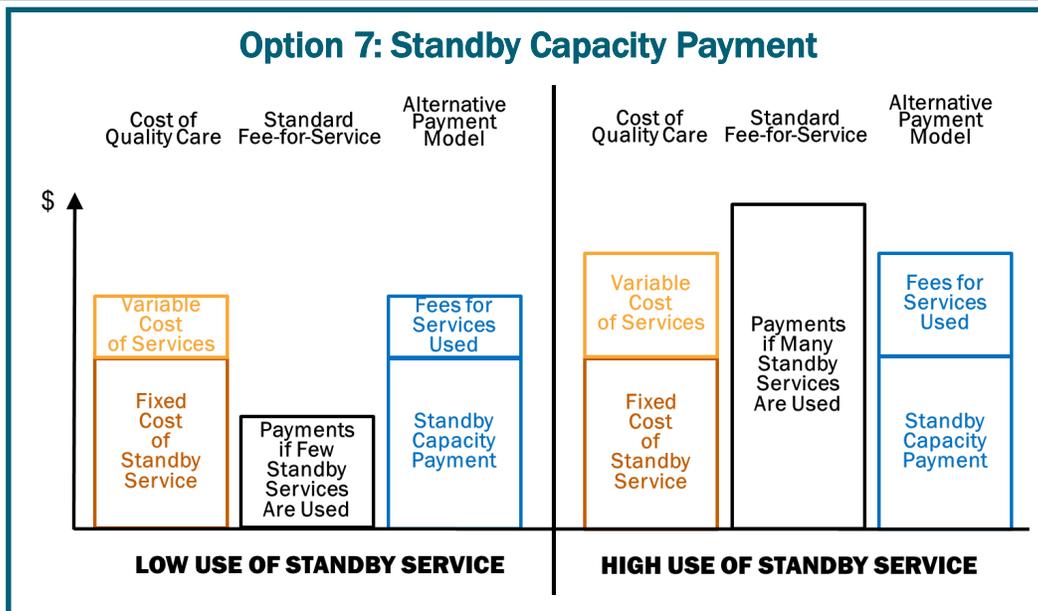
to the delivery of a specific number of services and may not even be tied to a specific type of service, so the period of time that the payment is intended to cover must be explicitly defined. Alternatively, the condition-based payment could define when or how often additional condition-based payments can be billed or paid for the same patient.

- For acute conditions, the condition-based payment could be based on the length of time needed for achieving a specific outcome. This could either be the time needed for full resolution or “cure” of the condition or it could be the time needed for the patient to advance to a subsequent phase of care. For example, whereas many commercial health plans pay for hospital care as a day-by-day service, Medicare pays a case rate for inpatient hospital admissions for medical conditions, and the case rate is expected to cover whatever length of time is needed for the patient to be safely discharged from the hospital, either to their home or to some form of post-acute care.⁸⁷
- For chronic conditions, if the services covered by the condition-based payment are expected to continue indefinitely, the period of time covered by the payment will inherently be arbitrary. One month is often selected as the length of a condition-based payment for a chronic condition in order to facilitate cash flow and to address changes needed because of changes in patient status (e.g., moving to another community, a change in the severity of the condition, etc.) or insurance coverage, but longer periods of time could also be used (e.g., three months), particularly for patients who have less severe conditions that require services less frequently.

Capitation and Population-Based Payment

Two special cases of Condition-Based Payment are what are typically labeled “practice capitation” and “global capitation.” Under *practice capitation*, a physician practice receives a pre-defined amount of money for each patient to support most or all of the services that the *practice delivers* to the patient for all of the patient's health conditions, but not for services the patient receives from other providers. Under *global capitation*, a physician practice or health system receives a pre-defined payment for *all of the services* the patient needs for *all* of their conditions. Such payments are typically paid for one month at a time (i.e., a “per-patient-per-month payment”). These types of capitation payments have recently come to be described as “population-based payments” because a provider does not receive them for individual patients, only for groups of patients.

In the past, capitation payments were not risk-adjusted or risk-stratified in any way, which meant that the revenues from the payments could differ significantly from the costs of the services the patients needed. Most capitation and population-based payments being used today are risk-adjusted or risk-stratified in some fashion based on the patient's health problems or other characteristics. This means that the payments are “condition-based,” but the patient does not have to have any particular condition or combination of conditions in order for the provider to be eligible to bill for the payment.



Option 7: Standby Capacity Payments

Rationale for Standby Capacity Payments

There are a number of important healthcare services for which a significant portion of the costs are not directly tied to either the number of patients who receive services or the number of services that patients receive. For example, a community will want to have a hospital Emergency Department adequately staffed and equipped to handle rare events such as serious accidents, natural disasters, infectious disease outbreaks, etc., while also hoping that no such events actually occur. It will also want both the Emergency Department and certain other hospital departments (such as a cardiac catheterization unit, a labor and delivery unit, a radiology department and laboratory, and a surgery suite) staffed and ready to quickly respond to heart attacks, strokes, premature births and complications of labor, major trauma, sepsis, etc. on a round-the-clock basis even if some of those situations occur only occasionally, particularly in small communities.

A service that must be ready to go on short notice is typically referred to as a “standby service” (i.e., the personnel and equipment needed for the service must be standing by in case a patient needs them). There are two kinds of standby services – patient-specific standby services and population-based standby services.

- **Patient-Specific Standby Services.** In some circumstances, a standby service may be associated with a specific patient; for example, if a patient is undergoing surgery and it is not clear whether an unusual procedure will be needed during the surgery, a specialist in that procedure may need to stand by so they can participate in the surgery immediately if needed. Medicare and other payers will pay physicians for the time they spend on standby for individual patients in specific circumstances.
- **Population-Based Standby Services.** In most cases, standby services are not associated with a specific

patient, but they are provided for the benefit of a large group or “population” of patients. For example:

- ♦ an emergency department maintains a certain minimum staff, equipment, and facilities to respond to a wide range of emergencies and other unexpected events in a timely way, but it is impossible to predict which individuals, if any, would potentially need to use emergency services. The population that benefits is the residents of the entire community in which the emergency department is located, as well as visitors to the community.
- ♦ a community needs at least one hospital with the ability to deliver babies even if every pregnant woman in the community would prefer to deliver their baby in a freestanding birth center, because complications will arise in some proportion of births in birth centers that will require transfer to a hospital for care. Although it is clear that this service will only be used by women in their childbearing years, it is impossible to predict how many of them will actually need to deliver their baby in the hospital.
- ♦ a community needs some inpatient hospital beds for patients with chronic diseases such as asthma, COPD, and heart failure who develop exacerbations of their disease that make them unable to breathe and require immediate treatment in a hospital emergency department but cannot be safely discharged the same day.

The cost of maintaining the minimum capability to serve an unknown number of patients can be described as “standby capacity cost.”

Although standby capacity is generally associated with hospitals because they are expected to be available to treat patients on a round-the-clock basis, there may also be a need to support standby capacity in certain types of physician practices and other types of providers. For example, it is important for people to have quick access to good primary care and certain specialists as well as hospital emergency services. Primary care practices, oncology practices, and others are increasingly being held accountable for having same-day appointments for patients in order to provide prompt treatment and discourage patients from using emergency departments for minor illnesses and injuries, so they are increasingly in-

curing standby capacity costs. Transportation services (both ambulance services and specialized non-emergency transportation, such as wheelchair vans for disabled individuals) must have slack capacity in order to be available when needed, and this is also a standby capacity cost.

Under fee-for-service payment systems, the only way a provider can cover its standby capacity cost is from the fees paid when services are actually delivered to a specific patient. For standby services that only need to be delivered occasionally, the price per service will likely be high because the average cost of each service will be based on the fixed cost of maintaining the capacity divided by the number of services actually delivered, not just the variable cost of delivering each additional service. The fixed cost will be higher if the service has to be available at all times than if the services could be scheduled in advance. For example, many hospitals charge high “trauma activation fees” when a trauma patient is treated; the high fee not only pays for the time devoted to that particular patient, but also for the time spent waiting in case a patient needs trauma services. In rural communities, higher fees are needed for standby services not because the cost of the standby capacity is higher, but because the volume of patients using the services is smaller, so the fee from each patient who actually receives a service has to be higher to add up to the same total amount needed to support the standby capacity.

Charging high fees to cover the costs of the service can be counterproductive if delivery of the service helps reduce the use of other, more expensive services. For example, if a primary care clinic in a rural area has to charge more per visit to cover its costs, patients will be less likely to use the clinic, and this could lead to delayed diagnoses and higher treatment costs. If transportation to the clinic is very expensive, patients will not be able to use it and will not be able to visit the clinic. The fewer patients who use the services, the less fee revenue there will be to support the services, creating a vicious cycle that could lead to loss of the service altogether. Many rural hospitals have closed because the amounts paid for their services are not high enough to cover the costs of those services and/or because the charges for the services are not affordable for patients without insurance.

If insurance is available and pays for each service based on average costs, the reverse problem can occur. The price will be far higher than the marginal cost per service (i.e., the additional cost incurred to deliver one more service), because so much of the cost associated with the service is the fixed cost of maintaining the ready-to-go capacity. This creates a financial incentive for the provider to deliver the standby service even when it is not necessary.

Structure of Standby Capacity Payments

Fee for service is not an appropriate way to pay for these “standby” services, because the services provide a benefit not just to patients who actually use them, but also to the individuals who could potentially need them. Instead, a “standby capacity payment” is needed from the “potential patients” as well as the patients who actually

receive services in order to support the standby capacity cost in a fair and adequate way. In order to create such a standby capacity payment, three things must be defined:

- **the beneficiaries of the standby capacity.** A definition is needed for the population of individuals who benefit from having the standby capacity available. For example, in the case of an emergency department in a rural community, this could be the residents of the community who would use the emergency department in case of an emergency, businesses whose employees would use it, etc. In the case of maternity care services, it could be all women of childbearing age in the community or all women who deliver a baby in any setting (i.e., at home or in a birth center as well as in a hospital).⁸⁸
- **the amount and cost of standby capacity.** A determination must be made regarding how much capacity should be maintained if the minimum number of patients used the service, and an estimate is needed for the cost or net loss that would be incurred in order to maintain capacity at the minimum volume of patients.
- **the formula for allocating cost among the beneficiaries.** The simplest approach would be to simply divide the cost by the number of potential beneficiaries. However, if there are differences in the level of benefit for different individuals, or if there are differences in the ability of beneficiaries to pay, then a more complex formula may be needed.

Example: In the case of a hospital emergency department, the majority of the potential patients are the residents of the community, so the residents of the community who would use the emergency department in case of an emergency could be asked to pay a standby capacity payment to support the emergency department. The total standby capacity cost could be defined as the cost the hospital would incur to have the bare minimum staffing available in the ED minus the revenue the hospital would receive at the lowest ED utilization it could expect to have. Each resident’s health insurance plan could then be asked to pay a fixed amount per month or per year determined by dividing the total standby cost by the total number of residents. If a business located in the community employs a large number of workers who do not live in the community or if it has a lot of out-of-town visitors, and if these workers and visitors would use the hospital emergency department in case of an accident, the business could pay for a portion of the standby capacity cost based on the number of workers it employs.

Example: In the case of maternity care services, the need for hospital-based labor and delivery services is limited to women who are pregnant, so a mechanism would be needed to identify those women and ask them or their health insurance plans to pay for the standby capacity cost of the hospital. For example, health insurance plans could pay the hospital a standby capacity payment for each woman who receives prenatal care or childbirth services paid for by the health insurance plan, including women who do not give birth in the hospital.

EXAMPLE OF HOW STANDBY CAPACITY PAYMENTS WOULD MATCH PAYMENTS TO COST

Assume that a community with 10,000 residents has a single community hospital with an emergency department (ED). As shown in Table 6a, the residents make a total of 3,000 visits per year to the ED and the hospital is paid an average of \$900 per visit. The Emergency Department costs \$2,575,000 per year to operate. 80% of the cost is fixed – the hospital incurs these costs whether it has any visits or not. With the current number of visits at current payment rates, the ED generates \$2,700,000 in revenue, creating a 5% profit margin for the hospital.

	CURRENT VISITS - FFS		
	\$/Visit	Visits	Total \$
ED Revenues			
Per Visit	\$900	3,000	\$2,700,000
ED Costs			
Fixed Costs (80%)			\$2,059,000
Variable Costs (15%)	\$172	3,000	\$516,000
Total Costs			\$2,575,000
ED Margin (5%)			\$125,000

Table 6b shows that a 10% reduction in the number of ED visits would cause the ED to lose money, because the cost would only decrease by 2%, but revenues would decrease by 10%. Conversely, as shown in Table 6c, a 10% increase in the number of ED visits would be highly profitable for the hospital because the cost of operating the ED would only increase by 2%, but revenues would increase by 10%. As a result, the hospital is financially harmed if better patient care reduces ED visits, and the hospital has a strong financial incentive to encourage greater use of the ED for non-emergency needs.

	CURRENT VISITS - FFS			REDUCTION IN VISITS – FFS			% Change
	\$/Visit	Visits	Total \$	\$/Visit	Visits	Total \$	
ED Revenues							
Per Visit	\$900	3,000	\$2,700,000	\$900	2,700	\$2,430,000	-10%
ED Costs							
Fixed Costs			\$2,059,000			\$2,059,000	
Variable Costs	\$172	3,000	\$516,000	\$172	2,700	\$464,400	
Total Costs			\$2,575,000			\$2,523,000	-2%
ED Margin			\$125,000			(\$93,400)	-175%

	CURRENT VISITS - FFS			INCREASE IN VISITS – FFS			% Change
	\$/Visit	Visits	Total \$	\$/Visit	Visits	Total \$	
ED Revenues							
Per Visit	\$900	3,000	\$2,700,000	\$900	3,300	\$2,970,000	+10%
ED Costs							
Fixed Costs			\$2,059,000			\$2,059,000	
Variable Costs	\$172	3,000	\$516,000	\$172	3,300	\$567,600	
Total Costs			\$2,575,000			\$2,626,600	-2%
ED Margin			\$125,000			\$343,400	+175%

EXAMPLE OF STANDBY CAPACITY PAYMENT (continued)

Table 6d shows how the ED could be paid using a Standby Capacity Payment. The insurance plan for each resident would make a fixed annual Standby Capacity Payment to the hospital of \$216 for that resident. If a resident came to the ED for diagnosis or treatment, the insurance plan would make an additional payment averaging \$180. (The actual payment would depend on the specific symptoms being diagnosed or the specific condition that was being treated.) If better chronic disease care or use of primary care for diagnosis and treatment of minor problems led to a 10% reduction in the number of ED visits, the hospital's revenues would only decrease by 2%, matching the 2% decrease in costs.

TABLE 6d	STANDBY CAPACITY PAYMENT			REDUCTION IN VISITS			% Change
	\$/Unit	Patients/ Visits	Total \$	\$/Unit	Patients/ Visits	Total \$	
ED Revenues							
Standby Capacity (Per Resident)	\$216	10,000	\$2,160,000	\$216	10,000	\$2,160,000	0%
Diagnosis/Treatment (Per Visit)	\$180	3,000	\$540,000	\$180	2,700	\$486,000	-10%
Total Revenues			\$2,700,000			\$2,646,000	-2%
ED Costs							
Fixed Costs			\$2,059,000			\$2,059,000	
Variable Costs	\$172	3,000	\$516,000	\$172	2,700	\$464,400	
Total Costs			\$2,575,000			\$2,523,000	-2%
ED Margin			\$125,000			\$122,600	-2%

As shown in Table 6e, if the hospital had a 10% increase in the number of visits, its revenues would only increase by 2%, matching the 2% increase in its costs. As a result, with the Standby Capacity Payment, the hospital would no longer have a financial incentive to increase ED visits, and it would no longer be financially penalized by efforts to reduce ED visits.

TABLE 6e	STANDBY CAPACITY PAYMENT			INCREASE IN VISITS			% Change
	\$/Unit	Patients/ Visits	Total \$	\$/Unit	Patients/ Visits	Total \$	
ED Revenues							
Standby Capacity (Per Resident)	\$216	10,000	\$2,160,000	\$216	10,000	\$2,160,000	0%
Diagnosis/Treatment (Per Visit)	\$180	3,000	\$540,000	\$180	3,300	\$594,000	+10%
Total Revenues			\$2,700,000			\$2,754,000	+2%
ED Costs							
Fixed Costs			\$2,059,000			\$2,059,000	
Variable Costs	\$172	3,000	\$516,000	\$172	2,700	\$567,600	
Total Costs			\$2,575,000			\$2,626,600	+2%
ED Margin			\$125,000			\$127,400	+2%

If a hospital or other provider receives standby capacity payments to cover the fixed cost of operating a service, there would still be a need for a fee or other types of payment when a service is actually delivered in order to cover the variable or semi-variable costs of the service. However, the amount of this per-service payment would be much lower than current fees for such services since it would be based on the marginal cost per service (i.e., the magnitude of the variable cost component), not based on the average cost of the service as fee-for-service payments typically are today. (A higher per-service fee could be charged to non-residents of the community so they would contribute to the fixed cost of the service, too.)

Examples of Standby Capacity Payments

Standby capacity payments are rare in healthcare. One notable example is that Medicare pays for a portion of the standby capacity costs in Critical Access Hospitals (small hospitals in rural areas that are located long distances from other hospitals) based on the proportion of patients who use the hospital's services who are Medicare beneficiaries. For example, the Critical Access Hospital can count as part of its eligible costs any payments it makes to physicians to be on-call to come to the emergency department during evenings and weekends, even if no patients actually come to the ED on an evening or weekend. If the only patients who used the hospital's services were Medicare patients, then Medicare would pay the full amount of those standby capacity costs, otherwise it would pay a percentage of the costs based on the proportion of the hospital's patients who were insured by Medicare.⁸⁹

Standby capacity payments are more commonly used to support services outside of health care.⁹⁰ For example, communities do not typically support their fire departments by charging fees to the victims of fires, nor do they support libraries by charging patrons when a book is borrowed.

Option 8: Volume-Based Adjustments

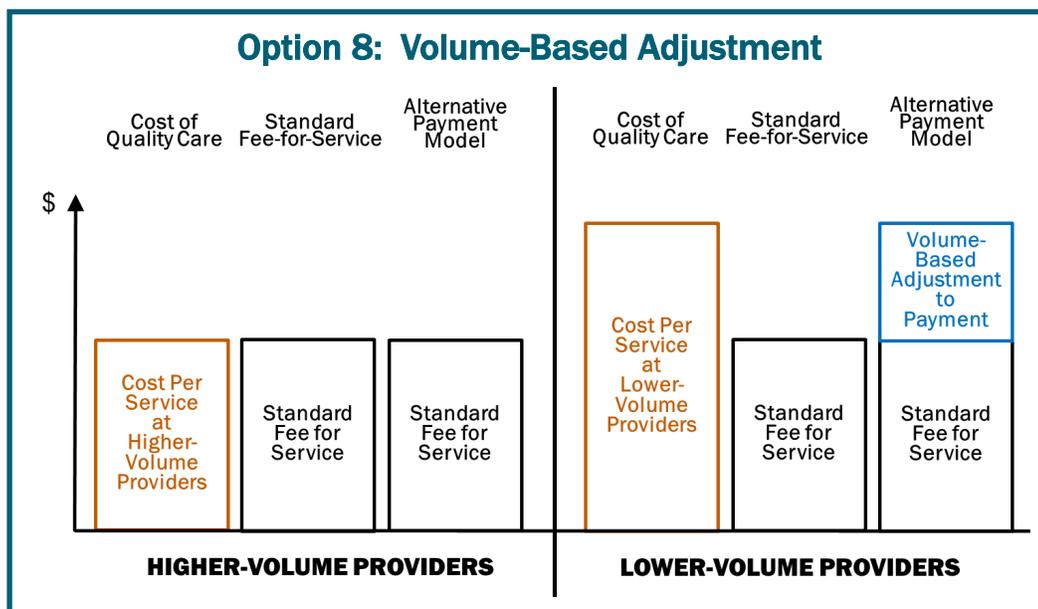
Limitations of the Standby Capacity Payment Approach

It is only feasible to use the standby capacity payment approach if there is a way to identify a specific population that (a) benefits from availability of the service over a period of time and (b) is willing and able to pay such a payment. For example, if the residents of a community are changing rapidly due to in-migration or out-migration, the residents or their insurance plans may resist paying a per-resident payment to support hospital standby capacity costs, since individuals may not expect to live in the community long enough to make the standby capacity payment seem like a good investment.

It is also difficult to use the standby capacity payment approach if there are multiple hospitals or other providers in a community that deliver standby services. For example, if there are two hospitals in a community that operate an Emergency Department, each of the hospital EDs will need to be on standby, but their combined standby capacity may be more than the total the community needs, and so a method of sharing standby capacity payments would be needed.

Rationale for Volume-Based Adjustments

An alternative approach when services have significant fixed costs is to pay on a per-service basis, but explicitly adjust the payment amount based on the total volume of the services delivered by the provider. Since the average cost per service will decrease if the volume of services increases, the payment amount per service could be reduced if the provider delivers a high volume of services and/or the payment amount could be increased if the provider delivers a low volume of services. This could help ensure that providers in small communities would be paid adequately for the higher average cost of services. It could also enable patients and payers to



benefit from the economies of scale that are often promised when providers consolidate.

A volume-based adjustment could reduce or eliminate the incentive to deliver unnecessary services that exists in a pure fee-for-service system when payment amounts are higher than marginal costs. It could also reduce or eliminate the financial penalty that providers would otherwise face when they eliminate unnecessary services by ensuring that fee revenues for the necessary services remain adequate to cover the costs of delivering services.

Examples of Volume-Based Adjustments

This approach has been used for adjusting payments to hospitals and other institutional providers in both the U.S. and other countries.

Example: In the Medicare program, a hospital can be paid more per discharge than other hospitals if it has fewer than 3,800 total discharges and if the hospital is located more than 15 miles from another hospital. The payment per discharge can be up to a maximum of 25% more than the standard DRG payment amount if the hospital has 500 or fewer total discharges.⁹¹

Example: The Medicare prospective payment system for end-stage renal disease (ESRD PPS) provides a 23.9% increase in payment for dialysis centers that have furnished less than 4,000 treatments in each of the 3 previous years.⁹²

Example: Maryland has an all-payer rate regulation system for hospitals. Although the state now sets an overall budget amount for each hospital, for many years it used DRGs as a method of controlling the costs of individual admissions and it used volume-based payment adjustments to control utilization. Originally, hospitals were only paid 50% of the standard payment amount for admissions above the projected volume level; if admission volume fell below the projected level, hospitals were also paid 50% of the standard payment amount for each admission that did not occur. In the early 1990s, the volume-based reductions were eliminated for the large teaching hospitals, and other hospitals received 85% of the standard amount when volumes increased. Between 2001 and 2008, the volume-based reductions were eliminated entirely, and then the 15% reduction (i.e., paying only 85% of the standard amount) was reintroduced in 2009. One analysis estimated that the removal of the volume-based reduction between 2001 and 2008 caused hospitals to spend 25% more than they would have otherwise.⁹³

Example: Germany pays hospitals a case rate for each admission, using a system of DRGs similar to what is used in the U.S. However, the hospital only receives 100% of the standard payment amount if the case mix-adjusted number of admissions is the same as the prior year. If the hospital and payers have agreed that the number of admissions will increase over the prior year, the hospital will only receive 75% of the standard payment amount for the

additional admissions. If the number of admissions increases even more, the hospital will only receive 35% of the standard payment amount for the additional admissions. Conversely, if the number of admissions decreases below the prior year level, the hospital can still be paid for 20% of the standard payment amount for the admissions that did not occur.⁹⁴

There are also examples of volume-based adjustments for physicians and other providers.

Example: Medicare and other payers use a "Multiple Procedure Payment Reduction" (MPPR) to reduce the payments for services when two or more services are performed on the same patient at the same time by the same provider. In most cases, the primary or highest-valued procedure is paid at 100% of the standard amount and then additional procedures or services are paid at 50% of the standard amount, but special rules apply to specific types of procedures.⁹⁵

Example: Other countries have "price-volume agreements" with pharmaceutical manufacturers that impose caps on the total amount the pharmaceutical manufacturer will be paid or that reduce the amounts paid as volume increases.⁹⁶

Structure of a Volume-Based Adjustment

Several decisions have to be made in order to use a volume-based adjustment:

- **Continuous vs. discrete adjustments.** Since most services will involve some amount of both fixed and semi-variable costs, the average cost per service will change even with a small change in the number of services. Matching these changes exactly would require setting a different payment amount for each discrete level of volume. An alternative is to define one or more specific thresholds at which payments will change and then keep payments the same within the range between a pair of thresholds.
- **The amount of adjustment.** The exact proportion of fixed costs may differ from provider to provider if there are choices about the method of delivering a service, about the supplier of a product, etc. Although some volume-based adjustment would be better than none, if a volume-based reduction for increased volume is too large, it could lead to shortages or underuse of desirable services.
- **Prospective vs. retrospective adjustment.** The actual total volume of services delivered by a provider and the cost of those services is only known after the services have already been delivered and likely after payments have already been made for most of the services. If a volume-based adjustment is needed, the adjustment could either be made retrospectively or prospectively:
 - ◆ A retrospective adjustment could be more accurate, but it would involve either refunding payments already collected or attempting to collect additional payments for patients who have already

received services, which would be administratively challenging.

- ◆ A prospective adjustment would involve charging more or less to future patients based on either historical cost and volume or a projection of future cost and volume.⁹⁷ The accuracy of this approach would depend on how consistent costs and the volume of services are over time. For example, if the volume of services decreases significantly over time due to ongoing reductions in avoidable services, then basing the payments on volume in the prior year could result in underpayments to providers who pursue such reductions more aggressively.

Differences Between Volume-Based Adjustment and Shared Savings

Although it might appear that a “shared savings” payment (similar to what is used in the Medicare Shared Savings Program) would achieve a similar goal, the percentage of the savings that is shared with the provider is typically an arbitrary amount that is not based on the amount by which the cost of a service changes when the volume changes. For example, if the fixed costs of delivering a service represent 70% of total costs at current volumes, then a 10% reduction in the volume of services would reduce the payer’s spending by 10%, but the provider’s cost would only decrease by 3%. A 50% shared savings payment would return 5% of the 10% reduction in spending to the provider, but that would still be 2% less than the new average cost of service. Conversely, if the volume of the service increased by 10%, the provider’s cost would only increase by 3%. If the provider was liable to repay the payer for 50% of the higher spending, it would still be financially advantageous for the provider to increase the number of services delivered. In contrast, a volume-based adjustment could be created that changes the payment for the service by 3% for each 1% change in the volume of services, thereby better matching payments to actual costs.

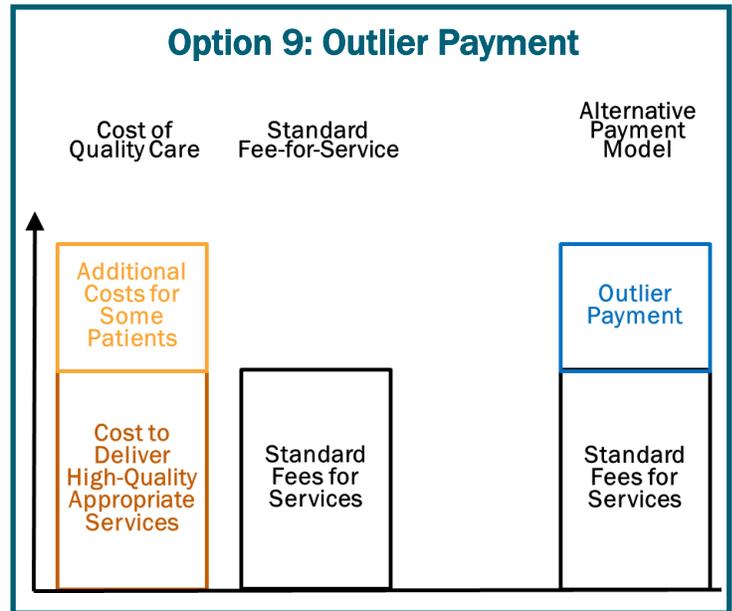
Option 9: Outlier Payments

The previous options all base payment amounts on the average amount of fixed, variable, or total costs across multiple patients. However, if there are *individual* patients who have unique characteristics that make the cost of delivering services dramatically higher than the average for the payment category to which they would be assigned, then a provider would be penalized financially if they have more such patients than other providers do. Moreover, if these patients exist, basing payment amounts on averages calculated across all patients will result in payments that are too high for most patients while still being too low for the outlier patients.

This problem can be addressed by (1) identifying individual patients whose services involve much-higher-than-average costs, (2) basing the standard payment amount on the average for the non-outlier patients, and (3) paying an additional “outlier payment” for the outlier patients.

Several current payment systems have a mechanism for making such “outlier” payments. For example:

- For most inpatient admissions, Medicare pays a standard amount (the “DRG payment”) for the entire hospital



stay and all of the services (other than physician services) that the hospital delivers as part of that stay. The payment amount depends on the patient's diagnosis and major procedure the patient received (as determined by the Diagnosis Related Group assigned to the hospital stay), but it does not depend on the length of stay or on the number or types of specific services that were delivered. However, if the cost of a specific patient exceeds the standard payment by more than a minimum amount, Medicare will pay 80% of the difference between (1) the cost of that patient’s stay and (2) the standard payment plus the minimum loss. (The cost of the stay is determined by multiplying the charges for the stay by the overall cost-to-charge ratio for the hospital, so it is only an approximation of the actual cost for the patient.)⁹⁸

- In the Medicare Physician Fee Schedule, there are special billing codes (CPT codes 99354-99357 “Prolonged Service With Direct Patient Contact”) that are used in addition to the usual Evaluation and Management code for a patient visit to indicate that additional time was spent with the patient or spent outside of the visit itself, and additional amounts are paid when these codes are billed. Modifier 22 (Increased Procedural Services) is added to the CPT code for a procedure to indicate that the work required to deliver the procedure was substantially greater than typically required. Medicare Administrative Contractors are permitted, but not required, to increase payments if there is adequate justification.

Structure of an Outlier Payment

Three decisions have to be made in order to define an outlier payment:

- **The types of situations in which higher-than-expected costs will be considered outliers.** There are several different types of situations that can cause costs to be higher than average for an individual patient and in which an outlier payment could be authorized:
 - ◆ The time or resources involved in delivering a particular service were higher than usual. For exam-

ple, a surgeon may experience unexpected challenges in performing surgery due to unique aspects of the patient's anatomy that significantly increase the length of the surgery.

- ◆ The number of services needed as part of a bundle of services were higher than usual. For example, a patient may take an unusually long time to recover from an illness, or a patient might need an unusually large number of visits with a physician or other healthcare professional in order to learn how to perform self-care properly.
- ◆ The amounts paid for products or supplies used in delivering services were higher than usual. For example, the patient's treatment requires use of a drug that is only produced by one manufacturer, and that manufacturer significantly increases the price of the drug.
- **The deviation threshold that will determine when an outlier case exists.** Since the standard payment for a patient will generally be designed based on an average cost for a range of patients, the mere fact that the actual cost for a patient is higher than the payment amount does not automatically justify receiving an outlier payment for the patient. An outlier payment is ordinarily triggered only when the cost or number of services needed by the individual patient is more than a pre-defined minimum deviation from the average. Using a smaller minimum deviation to trigger an outlier payment will mean the total payments will be closer to the provider's actual costs, but a smaller minimum deviation will also tend to reward inefficiency on the part of the provider.
- **The additional amount to be paid for the outlier case beyond what would otherwise be paid.** If a patient is identified as an outlier, one option is to make an additional payment equal to the difference between the actual cost for that patient and the standard payment amount. However, this also creates an incentive for inefficiency – once the costs are high enough to trigger an outlier payment, there is no incentive to control the additional costs. This can be addressed by paying only a portion of the additional costs, capping the amount of additional costs that will be paid, or defin-

ing an amount for the outlier payment that is not tied directly to actual costs.

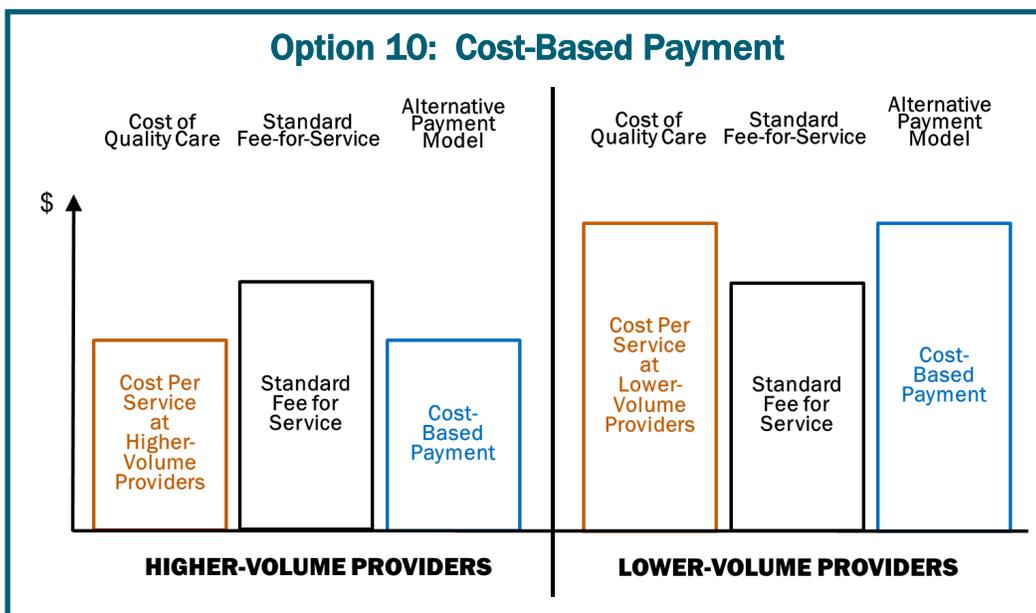
Stop-Loss Insurance

An alternative approach to outlier payments is to arrange for a separate insurance company to make the outlier payments by purchasing stop-loss insurance. The provider could pay the stop-loss insurance company a premium payment for each patient who receives the service for which the outliers may occur, and then the stop-loss insurer would make an additional payment to the provider when an outlier does occur. The stop-loss insurance policy would need to specify each of the three items described above – the situations covered by the policy, the deviation threshold that triggers an insurance payment, and the amount that would be paid when that situation occurs and the threshold is met. The payments to the provider for services would need to be large enough to pay for these premiums. If a provider is large enough or has adequate financial reserves, it could set aside a portion of the payments so that it could self-insure when outliers occur.

Option 10: Cost-Based Payments

Options 1-8 are all “prospective payments,” i.e., the payment amounts are set prior to a service being delivered. However, this also means the payment amounts may or may not match the actual costs that a provider incurs in delivering a service or combination of services. For example, hospitals in isolated rural communities often have difficulty attracting physicians, nurses, laboratory technicians, etc. and not only do they have to pay much higher amounts to attract and retain them than hospitals located in metropolitan areas, the amounts they have to pay may vary significantly and unpredictably from year to year.

An alternative to a prospective payment is a “cost-based payment” that explicitly ties the payment amount to the actual cost a provider incurs for delivering a service or combination of services. Although cost-based payment



is no longer used as widely as it once was, it is still used in the Medicare program for specific kinds of providers:

- **Critical Access Hospitals:** Medicare pays small, rural hospitals designated as Critical Access Hospitals based on the actual costs incurred in patient care rather than pre-defined rates. The amount that the hospital spent during the year on allowable costs is divided by the total charges for patients who received services in order to determine a cost-to-charge ratio. The charges for Medicare patients are multiplied by the cost-to-charge ratio and then by 101% to determine the amount that Medicare will pay. However, under federal sequestration rules, the hospital's payment must be reduced by 2%, so the hospital is only paid 99% of its allowable costs, and since not all costs are allowable (for example, the time physicians spend with patients is not an allowable cost), the effective rate is lower than 99%.⁹⁹
- **Rural Health Clinics:** Medicare pays Rural Health Clinics using a cost-based methodology similar to Critical Access Hospitals, except that there is also a maximum amount per visit that Medicare will pay.¹⁰⁰
- **Cancer Hospitals:** Medicare pays 11 hospitals that primarily treat cancer patients ("cancer hospitals") for their inpatient services based on their actual costs, limited by a hospital-specific maximum amount. Each cancer hospital (referred to as a PPS-Exempt Cancer Hospital, or PCH) has a cost ceiling based on the average cost per discharge that it incurred during a base year plus an annual update for inflation. Medicare pays the lower of the hospital's actual costs or the ceiling, but if the costs exceed 110% of the ceiling, the hospital also receives a "relief payment" equal to the lesser of (a) 50% of the inpatient operating costs in excess of 110% of the ceiling and (b) 10% of the ceiling. If the costs are less than the ceiling, the hospital receives a bonus payment equal to the lesser of (a) 15% of the difference between actual inpatient costs and the ceiling and (b) 2% of the ceiling.¹⁰¹

Cost-based payment is generally viewed as less desirable than prospective payment options because it does not limit how high the payment for a service can be. However, unlike other options, cost-based payment can limit how much the *profit* on a service can be. Under fee-for-service payment, the average cost of a service will decrease if the service is delivered more frequently and that will lead to a higher profit if the payment per service is fixed. In contrast, under cost-based payment, the payment would decrease if the average cost per service decreased, and so the level of profit would remain the same.

Although cost-based payment does not encourage higher costs *per se*, neither does it create an incentive for reducing costs. A provider that incurs higher costs will not receive significantly greater profits but neither will the provider experience a financial loss.

Cost-based payment can be desirable if delivery of a new service is expected to result in significant savings on other services but there is either (a) considerable uncertainty about exactly what the new service will cost or (b) considerable variation in the cost of delivering the service in different communities or settings. In these cases,

paying based on the actual cost may be desirable or even necessary for encouraging delivery of the new service, at least during the initial phases of implementing an APM until providers are comfortable that they understand the magnitude of the costs and can be comfortable accepting a specific pre-defined payment amount.

Option 11: Using Multi-Component Payment Structures

Why "Simple" Payment Systems Are Unlikely to Align With the Cost of Care Delivery

Options 1-9 are each designed to align payment with one aspect of costs – either fixed costs, semi-variable costs, or variable costs – but not with all three. Since most services involve a combination of fixed costs, semi-variable costs, and truly variable costs, none of the options is ideal for matching payment to costs at different volumes of services.

For example, a condition-based payment system (in which a physician practice, hospital, or other provider receives a fixed payment for a patient with a particular condition or combination of conditions) does a better job of matching a provider's fixed and semi-variable costs than paying a fixed fee for each service, but it does a much worse job of matching the provider's variable costs. As a result, although the provider would no longer experience higher profits when unnecessary services are delivered, the provider's profits now increase when *necessary* services are *withheld*, which creates concerns about undertreatment of patients.

The mismatch between payment and costs is a key reason why most current value-based payment systems either fail to solve the problems with fee-for-service payment or create different kinds of problems for patients, payers, or providers. "Simple" payment systems based on any one of the options cannot be expected to match a provider's costs for delivering a service because those costs are not generated through an equally simple approach.

In fact, in most cases where seemingly simple payment options have actually been implemented, complex processes are included to try and mitigate the risks of underpayment or overpayment created by the basic payment model. For example, most capitation payments to provider organizations are not truly "global" payments that cover all services their patients receive. Although the provider organization receives a monthly payment for each assigned patient, there is also a Division of Financial Responsibility that contains a lengthy and complex list of services that the capitated organization is not expected to pay for using the monthly payment and that the insurance plan will pay for directly using fee for service payment or some other method.¹⁰² These systems will generally increase administrative costs and create a complex set of interactions among the different incentives, when the real goal is to try and better align payment with costs.

Using Multi-Component Payment Models to Align With Service Delivery Costs

In order to ensure payment amounts are adequate to cover providers' costs while also encouraging providers to control their costs, a payment model can be created that explicitly includes separate components using two or more options from Options 1-10. For example, a payment model might include all four of the following components:

- a standby capacity payment (Option 7) to address fixed costs that must be supported regardless of the number of patients or services; and
- a condition-based payment per patient (Option 6) to support semi-variable costs, such as personnel, that only vary when there are large differences in the number and types of patients served; and
- a fee per service payment (Option 1) to support the variable costs that differ significantly even if one more or one fewer service is delivered; and
- an outlier payment (Option 9) for the small subset of patients who require unusually large amounts of time, specialized services, or expensive supplies.

Each of these payment components would be paying for a different portion of the total cost of delivering services. In a typical fee-for-service payment system, all of these different portions of the cost would be added together and the total cost would be divided by the number of services delivered to determine an overall average payment amount for the service. The provider's total revenue for the service would then equal the number of services delivered times that payment amount. In a multi-component payment system, a separate payment amount would be determined for each aspect of cost, and the provider's total revenue for the service would then equal the sum of all of the individual component payments.

Example: Assume that a physician practice treats patients who have a particular health problem (such as cancer) using expensive drugs administered to the patient in the practice's offices. The total cost for the practice to treat the patients will include the amount it spends for the pharmacists who purchase and store the drugs, the salaries and benefits for the nurses who perform the drug infusions, and the amount the practice pays to purchase individual drugs for individual patients.

- *The cost of the pharmacists and the pharmacy equipment is essentially fixed – it will be the same whether one patient per week receives treatment or dozens of patients do. In a small community, a standby capacity payment may be needed to ensure the practice can maintain the capacity to provide treatments even if no patients happen to need treatment at particular points in time.*
- *The number of nurses needed will depend on how many patients are likely to be treated, but the number will not change if an individual patient receives one more or one fewer treatment. A condition-based payment could be used to support the nursing staff, so if more patients are treated, there will be adequate revenue to enable additional nurses to be hired.*

- *The cost of purchasing the drugs will vary tremendously depending on the number of drugs administered and the prices of the drugs, so a per-service payment could be used to pay for the drugs themselves.*

3. Enabling Control of Services Delivered by Other Providers

As explained in Section V, a different type of barrier in the current payment system is created by the fact that payments for individual services are made separately and independently of each other. If the opportunity for savings requires that two or more providers each work as a team to deliver certain services in a particular way, a mechanism is needed to ensure that each provider completes their portion of the overall approach to care delivery appropriately.

Option 12: Multi-Provider Bundled Payment

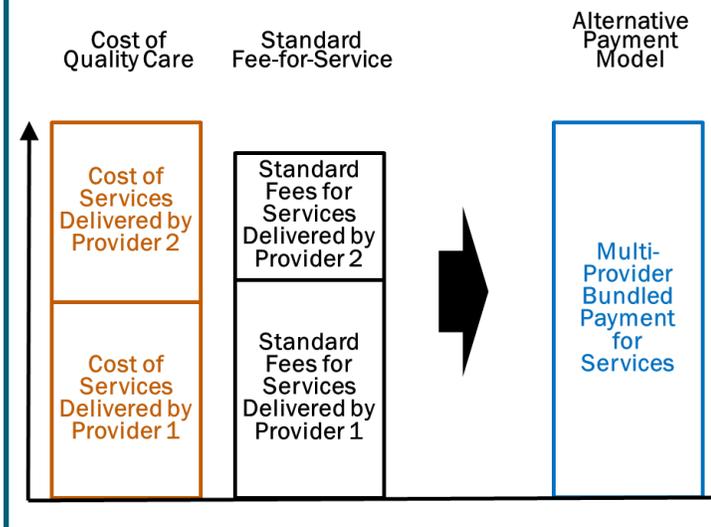
One solution is to pay a “multi-provider bundled payment” that supports all of the individual services delivered by all of the providers on the team that is delivering care to the patient. Since multiple providers would have to receive portions of the revenue to cover their costs, the payment would need to be made either to (a) one of the providers, who would then pay the other providers for their services when they are delivered in the way needed to achieve the planned efficiencies, or (b) an organized partnership of the providers that receives the bundled payment and divides it among the individual provider members.

Example: In order to support integrating behavioral health services in primary care, primary care practices can bill Medicare for “Psychiatric Collaborative Care Services” (CPT Codes 99492, 99493, and 99494). The primary care practice receives a monthly payment to support services delivered by the primary care physician or clinician and a behavioral health care manager, and a portion of the payment is paid to a psychiatrist who consults with the PCP and behavioral health care manager to determine the most appropriate treatment plan for the patient.¹⁰³

Example: In the Medicare Acute Care Episode Demonstration, a participating hospital and surgeon were not paid separately for their services as part of an orthopedic or cardiac surgery procedure; instead, a single bundled payment was paid. The participating hospitals and physicians were required to have or create a Physician-Hospital Organization (PHO), and the bundled payments were paid to the PHO. The PHO was then responsible for dividing the payments between the hospital and the physicians.¹⁰⁴

The purpose of bundling here includes but expands upon the goals described in Option 2. As with a bundled payment to a single provider, the multi-provider bundled payment provides flexibility regarding the services that

Option 12: Multi-Provider Bundled Payment



can be delivered as well as certainty for the payer as to how much will be spent. However, a multi-provider bundle also (1) creates flexibility as to which provider will deliver services and (2) gives the providers who are involved a way of achieving certainty about what services the others will be delivering and the total cost that will be incurred. For example, a bundled payment that includes both the payment for a hospital stay and for post-acute care services would enable the hospital and post-acute care providers to mutually agree on how much of the care would be delivered in the hospital and how much and what types of post-acute care services would be used (e.g., services in a skilled nursing facility, from a home health agency, from an independent physical therapy practice, or some other type of provider).

As with the single-provider bundled payments discussed earlier, there will need to be definitions for:

- the services included in the bundled payment;
- the length of time covered by the bundle; and
- the “trigger” for the payment (i.e., what service, condition, etc. will make the provider eligible to receive the payment).

The amount of the bundled payment will also need to be determined, and it will need to be adequate to cover the costs of the services it is designed to support. If different patients need different numbers and types of the services included in the bundle, then a single bundled payment will be problematic because it could be higher

than what would have otherwise been spent or lower than necessary to cover costs. Addressing this will likely require stratifying the payment as discussed in Option 5.¹⁰⁵

Multi-provider bundled payments can be triggered either by the delivery of one or more specific services in the bundle (e.g., a hospital procedure) or by the existence of a specific health condition. Condition-based multi-provider bundles provide the greatest flexibility as to which services could be delivered through the bundle and who can deliver them, but determining how to allocate the bundle and who should be involved in making that decision may be more challenging because of the broader range of services and providers that could be involved. In addition, if there are significant fixed costs associated with one or several of the services, then it may make sense to pay for the bundle using a multi-part structure such as described in Option 11.

Bundles Do Not Automatically Result in More Coordinated Care

A multi-provider bundled payment works best when the providers have agreed to work as a team and the patient has agreed to receive all of the services in the bundle from the members of that team. If the bundled payment is not adequate to cover the costs of delivering all of the services in a high-quality way, or if a high-quality provider is not assured of receiving an adequate share of the bundled payment, the patient may receive lower-quality care than if the services were paid for separately. Moreover, if the patient is unwilling to use only the members of the team for services and the bundled payment does not require them to do so, payments made to the non-participating providers could result in financial penalties for the providers on the team when the reconciliation process occurs.

In addition, as with single-provider bundles, bundled payments can be problematic for patients who need more of the individual services contained in the bundle or who need more expensive versions of the services. These patients could receive fewer services than they need, particularly if the bundled payment is not stratified or risk-adjusted based on differences in patient needs.

Example: Most current APMs with bundled payments combine the payment for a hospitalization with payments for post-acute care, and most of the savings that have been achieved through these payments have resulted from the use of fewer and less expensive post-acute care services.¹⁰⁶ However, this can also result in inadequate care for patients who need more intensive post-acute care services.¹⁰⁷

4. Modifying Cost-Sharing

If the amount that a patient is required to pay in cost-sharing for a desirable service could discourage or prevent patients from using it, then the cost-sharing amount may need to be reduced. Conversely, if the amount the patient is required to pay for a low-value service is so low that it is encouraging them to use the service inappropriately or is making it difficult for a provider to substitute a higher-value service, then cost-sharing for the service may need to be increased.

The two options described below can be used in conjunction with Options 1-12, or they can be used by themselves if patient cost-sharing is the only barrier to implementing different services and achieving savings opportunities.

Option 13: Modify standard cost-sharing rules

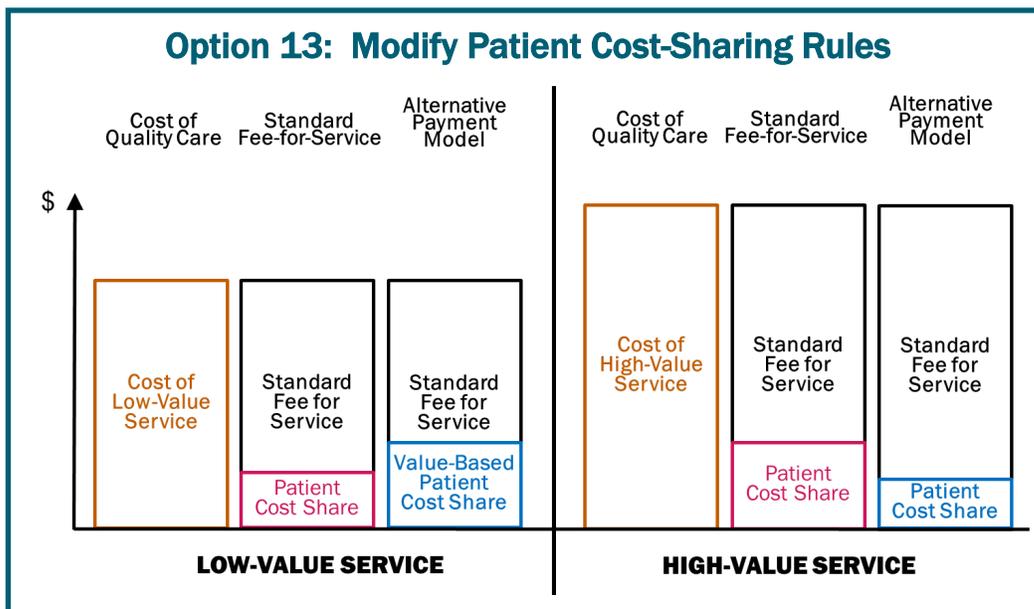
In most insurance plans, the amount that a patient is expected to pay for a healthcare service is determined using some combination of the following approaches:

- The service (or the provider delivering that service) may not be covered by the health insurance plan under any circumstances, in which case the patient would need to pay the full price of the service;
- If the service is covered, there may be a deductible that has to be met (i.e., a minimum total amount the patient must spend on covered healthcare services) before the insurance plan contributes anything to pay for the service, which also means the patient could have to pay as much as the full price of the service;
- There will likely be a copayment or co-insurance amount the patient must pay for the service after any deductible is met; and
- There may be an out-of-pocket limit (OOP) on the total amount of cost-sharing the patient is required to pay during a particular period of time.

Cost-sharing barriers often arise in alternative payment models when standard formulas for the components described above are used to determine the cost-sharing amounts for new or existing services related to the APM, but the resulting cost-sharing amounts for the patients are not consistent with the goals of the APM. For example, in the Medicare program, beneficiaries are required to pay 20% cost-sharing for all outpatient services except for preventive care services. If the APM allows a physician practice to bill for a new service that is expected to reduce avoidable hospitalizations, patients may be unwilling to pay 20% of the cost of the new service if they don't believe they are really at risk of hospitalizations. However, even though it would cost Medicare more to reduce the cost-sharing, this might be offset by the savings from fewer hospitalizations if the patients use the service and it is effective.

The obvious solution to this is to define special cost-sharing requirements for one or more of the services delivered under the APM. For example, cost-sharing for a desirable service could be reduced by exempting the service from the deductible and/or by reducing the size of the co-insurance or copayment requirement. If the goal of the APM is reduce utilization of a less effective service, the copayment or co-insurance for that service could be increased.

Example: In the "Initiative to Reduce Avoidable Hospitalizations Among Nursing Facilities - Payment Reform," CMS makes additional payments to both Skilled Nursing Facilities and to physicians (or other clinicians) for treatment of specific types of health conditions, but the Medicare beneficiaries receiving the services are not charged any co-insurance or deductible amount for these services.¹⁰⁸



Option 14: Create or change last-dollar cost-sharing amounts

The standard cost-sharing requirements described above all represent “first dollar” cost-sharing approaches, i.e., the amount that the patient pays is determined first, and then the payer pays the rest. For example, if an insurance plan includes a deductible, the patient has to spend enough in total on health care services to meet the deductible before the insurance plan shares in the cost of any service. If there is an out-of-pocket limit on the patient’s cost-sharing, once the patient spends enough on services to reach that limit, the insurance plan begins paying the entire amount.

As noted in Section V, this approach creates perverse incentives for the patient because once the total payment amount for a service exceeds the limit on the patient’s cost-sharing (based on the combination of the copayment, co-insurance, deductible, and/or out-of-pocket limit), the cost to the patient will be the same even though the total cost for the insurance plan may be very different. Option 14 would require the patient to pay the “last dollar” of the cost, i.e., if there are two different choices of services or providers, the patient’s cost sharing would be based on the *difference* in the cost or total payment amount.

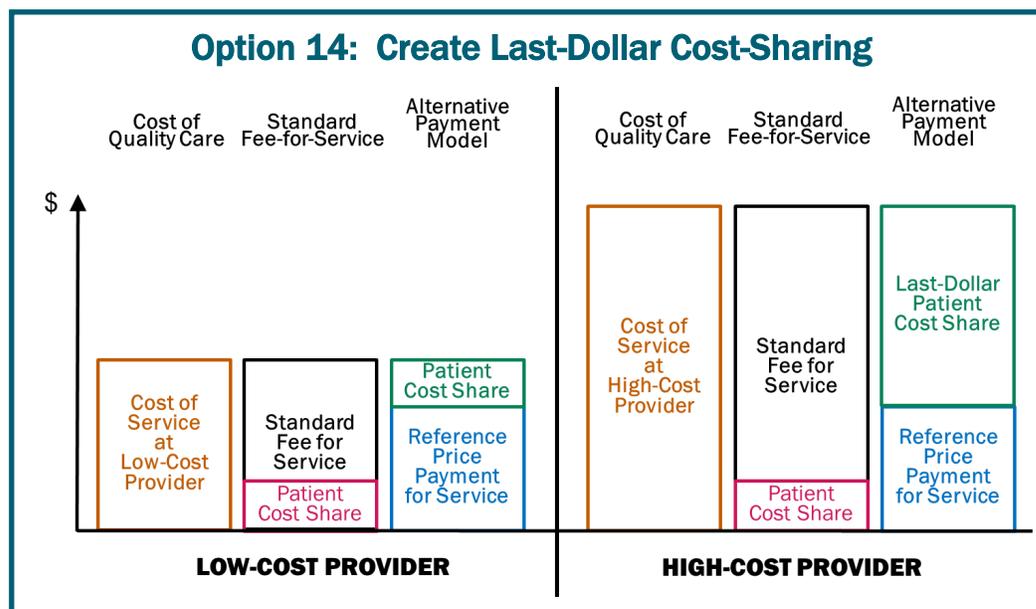
Example: Assume that a patient needs knee surgery and there are several provider teams that deliver the procedure. Assume further that the amounts the providers charge for all of the services associated with surgery (the surgeon’s work, the hospital stay, and any post-acute care needed) vary from \$20,000 to \$60,000. Assume that the patient’s insurance plan has a \$5,000 deductible and requires the patient to pay 20% of the cost of the service up to a maximum out-of-pocket limit of \$10,000. Then in order to receive surgery from the \$20,000 provider, the patient would have to pay \$10,000 (the \$5,000 deductible plus 20% of the remaining \$35,000 would equal \$12,000, so the patient would only be required to pay \$10,000 since that is the out-of-

pocket limit), and in order to receive surgery from the \$60,000 provider, the patient would also have to pay \$10,000 (the \$5,000 deductible plus 20% of the remaining \$55,000 would equal \$16,000, which is also more than the out-of-pocket limit). The result is that the patient would see no difference in cost-sharing between the two providers, even though the insurance plan would have to pay much more (\$40,000) if the patient uses the second provider.

Instead, the patient could be required to pay \$10,000 for surgery from the first provider and up to \$50,000 for surgery from the second provider based on the \$40,000 difference in the prices of the two providers.

There are at least two different ways to implement last-dollar cost-sharing:

- **Balance billing.** The payer could set a specific amount it is willing to pay for the service, and then the patient would be required to pay any additional amount if they use a provider that charges more for that service. In the knee surgery example, the payer could agree to pay \$20,000 for knee surgery wherever it is performed, and the patient would pay nothing to the first provider and \$40,000 to the second provider.¹⁰⁹
- **Reference pricing.** The payer would set a maximum amount it is willing to pay based on the distribution of amounts that different providers charge, and then the patient pays the difference between the reference price and the amount their provider charges. For example, if the payer sets the reference price at the 60th percentile of provider charges, then if 60% of knee surgery providers charge \$45,000 or less, the payer would set the “reference price” at \$45,000. If the patient chooses a provider that charges less than \$45,000, the patient would pay whatever standard cost-sharing amount is required, but if the patient chooses a \$50,000 provider, the patient would pay an additional \$5,000 (the difference between the \$50,000 charge and the reference price).¹¹⁰



**TABLE 7
OPTIONS FOR REMOVING THE BARRIERS IN THE CURRENT PAYMENT SYSTEM**

Payment Option	Payment Barrier(s) Addressed	Challenges/Weaknesses
1. Pay a fee for the service	No payment for a high-value service	Can encourage unnecessary use
2. Bundled payment for a group of services	No payment for a service that complements or substitutes for other services	Can limit flexibility if patients need different combinations of services
3. Higher payment for the service	Payment is usually below cost	Can encourage unnecessary use
4. Payment stratified by phase of care	Payment too low in some phases	Requires clear definition of phases
5. Payment stratified by patient characteristics	Higher cost of delivering service to certain types of patients	Requires objective way of assessing presence of characteristics
6. Condition-based payment	Cost depends more on number and type of patients than # of services	Can encourage over-diagnosis of condition
7. Standby capacity payment	Service needs to be available even if no patients need or use it	Requires determining minimum capacity needed for service
8. Volume-based payment adjustment	Higher cost for low-volume providers	Can encourage delivery of low volumes of service
9. Outlier payment	Higher cost for specific patients	Can reward inefficiency
10. Cost-based payment	Costs differ for different providers	Can encourage inefficiency
11. Multi-component payment	Cost of services depends on multiple factors	Increases the complexity of payment
12. Multi-provider bundled payment	Multiple providers need to deliver services in a coordinated way	Requires designating a payment recipient and allocation method
13. Modified first dollar cost-sharing	Co-pays, co-insurance, deductibles discourage use of high-value service	Lower cost-sharing can encourage unnecessary use
14. Last-dollar cost-sharing	Different providers/services have similar benefits but different costs	Can discourage use of higher-cost services that have better outcomes

B. APM Component #2: Creating Accountability for Spending

If Component #1 of the Alternative Payment Model eliminates or adequately mitigates the payment barriers identified in Section V, then it should be feasible for patients to receive a more desirable combination of services such as those defined in Section IV. However, before implementing these changes in payment, a payer or patient will want assurance that the expected savings will actually materialize, or that improvements in quality have not been accompanied by increases in spending. Consequently, in addition to Component #1, an Alternative Payment Model requires one or more components that create accountability for spending.

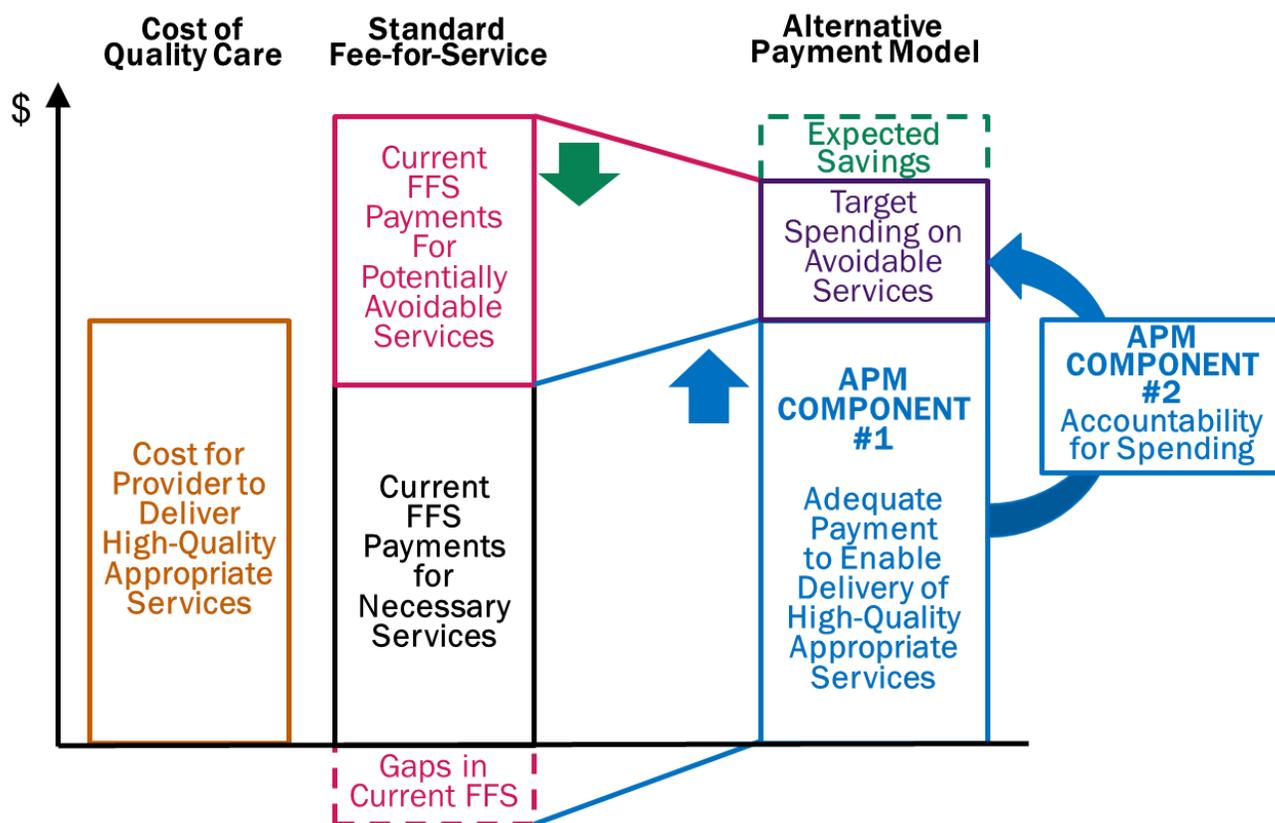
An accountability component for spending has four distinct elements:

1. One or more *measures of spending or utilization* that the participants in the APM will be accountable for reducing or controlling;
2. A *Target* for each of these measures, i.e., the level that must be achieved or maintained or the change that must occur in order for the APM to be deemed successful in achieving its goal;

3. A *performance assessment methodology*, i.e., the calculations that will be made to determine whether a specific entity participating in the APM has achieved or maintained the targets.
4. A *mechanism for adjusting payments based on performance*, i.e., what changes will be made in payments if the targets are not achieved.

It will often be desirable to have multiple accountability components for different aspects of spending. For example, if the APM is intended to achieve savings through reductions in specific services, a payer would want to ensure those reductions were achieved, but the payer would also want to ensure there were not increases in other related services that offset the savings from the targeted services. To address this, one accountability component could be designed to ensure savings in the targeted services, while one or more additional components could be designed to ensure there were not significant increases in spending on other related services.

FIGURE 2
APM COMPONENTS #1 AND #2
ALLOW ADEQUATE PAYMENT WITH LOWER SPENDING



1. Defining the Accountability Measures

An accountability component must first define the specific aspects of utilization or spending for which the participant in the APM will be accountable and how they will be measured.

a. Types of Spending Affected by the APM

If the APM is explicitly intended to reduce or control spending on certain types of services, then the APM needs to have specific measures for each of those services or the aspects of spending which are to be reduced. This could include measures of:

1. **Planned reductions in utilization or spending on services delivered by the APM participants.**
2. **Planned reductions in utilization or spending on services ordered from other providers.**
3. **Reductions in utilization or spending on unplanned services that the APM is intended to achieve.**

Different APMs will focus on different types and subsets of services, so a measure of spending that is appropriate for one APM may be too broad or too narrow for another APM. If the patient could receive the same type of service under circumstances that are unrelated to the APM, the spending measure will need to be defined precisely enough to only include the circumstances targeted by the APM.¹¹¹

In addition, the APM may need definitions/measures of other aspects of spending that could be affected by the APM, since increases in those areas could offset any savings from the specific types of spending targeted by the APM. There are at least four additional categories of spending that should be considered:

4. **Spending on Complications of Treatment.** If the APM supports delivery of a new service or expanded delivery of an existing service, and if there is the potential for patients to experience complications or other adverse effects from that service, higher spending to treat any such complications would offset the savings from the desired effects of the service, so the APM would need to measure the rates of those complications and the costs of treating them.
5. **Spending on Complications of Undertreatment.** If the goal of the APM is to reduce the use of an existing service, and if patients could experience complications or other adverse effects by *not* receiving that service, the spending to treat the complications would offset the savings from not delivering the service, and so the APM would need to measure the rate of those complications and the costs of treating them.
6. **Spending on Substitutions of Other Services.** If the goal of the APM is to reduce the use of a particular service, but there are alternative services that could be substituted for that service, then an increase in use of the alternatives would offset the savings from reduced use of the targeted service. Consequently, the APM would need to measure utilization or spending on the substitute services as well as the spending on the services the APM is expected to reduce. For example, if the goal of the APM is to reduce utilization

of MRIs for lower back pain, the APM could also measure the use of other types of imaging studies, in addition to measuring the number of MRIs, in order to avoid the possibility that a provider is substituting CT scans for MRIs. If the goal of the APM is to reduce hospital readmissions, it may be appropriate to measure observation stays as well as actual admissions to avoid the possibility that patients are still receiving hospital care but it is being classified differently.

7. **Spending from Increased Utilization of a Lower-Priced Service.** If the goal of the APM is to reduce the cost of delivering a particular service, but patients and/or providers have discretion about whether to use the service, an increase in utilization of the service would offset the savings achieved from delivering each service at a lower cost. Consequently, the APM would need to measure overall utilization and spending on the service as well as the amount of spending each time the service is used. For example, if an APM that is designed to reduce the total cost of knee replacement procedures results in an increase in the number of knee replacement procedures, there would be less savings than what would have been expected based on the change in the cost of each procedure alone, and it is possible there would be no savings at all.¹¹²

If the APM could have two or more of these effects, then definitions/measures of the spending related to each effect will be needed.

b. Composite Measures vs. Service-Specific Measures

Problems with Total Cost of Care Measures

Rather than defining and measuring spending for specific types of services, many APMs have used a “total cost of care” measure for the patients receiving services supported by the APM. A total cost of care measure includes spending on *all* types of services the patient receives, including not only the services targeted by the APM or related to the APM, but also services that are unlikely to be affected in any way by the APM.

From the payer’s perspective, a total cost of care methodology has the advantage of simplicity and comprehensiveness. There is no need to precisely define and measure spending on the specific types of services where reductions are expected or where increases may occur, nor is there any concern that spending is being shifted to unmeasured services; the payer merely needs to measure the total spending on the patients to determine if savings have occurred.

What is simpler for payers, however, can be very problematic for the providers who are being held accountable under the APM and for the patients who are receiving services supported by the APM. The total cost of care for patients can increase or decrease for reasons that are completely unrelated to the services supported by the APM, and this could result in an inappropriate

penalty or reward for the APM provider. In addition, placing providers at financial risk for the total cost of care when they cannot control all aspects of care can penalize them for serving high-need patients and create undesirable incentives to stint on needed services. If adjustments are made to the APM to try and address these problems, such as adding more quality measures or more elaborate risk adjustment systems, the resulting methodology will no longer be “simple.”

Example: The Oncology Care Model (OCM) is an APM created by the Center for Medicare and Medicaid Innovation (CMMI) for oncologists treating cancer patients with chemotherapy. Oncologists are eligible to receive a performance-based payment based on Medicare spending during the six months following initiation of chemotherapy. The spending measure used in OCM includes spending on all types of services that the patient receives, including (a) spending on services unrelated to their cancer that the oncologist has no ability to control and (b) spending on cancer drugs even though neither CMS nor the oncologist can control the prices of those drugs.¹¹³ As a result, an oncology practice might successfully reduce unnecessary utilization of cancer-related services for its patients but see total spending increase because physicians in other specialties used more services (or more expensive services) to treat the patients’ non-cancer-related problems or because of increases in the prices of the drugs used to treat the patients.

Strengths and Weaknesses of Episode Spending and Other Composite Measures

An alternative to a total cost of care measure is a more narrowly-defined composite measure that includes only services related to the specific condition for which the patient is being treated or to a specific procedure the patient has received. For example, “episode spending” measures for surgeries typically are defined to include spending on (a) the hospitalization when the surgery was performed, (b) rehabilitation services the patient receives after discharge, (c) any rehospitalizations that occur within 30-90 days after discharge, and (d) other services such as physician visits, tests, etc. related to the patient’s surgery. If an APM is designed to improve care for patients with a chronic disease, then a “condition-based spending” measure could be defined based on all of the services patients receive that are related to their chronic disease (including office visits with their primary care physician and specialists, medications to treat their disease, hospitalizations for exacerbation of the disease, etc.).¹¹⁴

The total cost of care for patients can increase or decrease for reasons completely unrelated to the services supported by the APM, and this could result in an inappropriate penalty or reward for the APM provider. In addition, placing providers at financial risk for the total cost of care when they cannot control all aspects of care can penalize them for serving high-need patients and create undesirable incentives to stint on needed services.

Episode spending and condition-based spending measures are much more difficult to calculate than a total cost of care measure because judgments have to be made about which services will be included and whether the circumstance in which a particular service was delivered is truly related to the episode or condition. Although there are episode “grouper” software packages designed to automate these determinations, they can erroneously include unrelated services and exclude related services, particularly when the algorithms are based only on information derived from claims data.¹¹⁵

Moreover, the fact that a service is clinically related to a procedure supported by the APM or a condition being addressed by the APM does not mean that those involved in the APM should be held accountable for all changes in the utilization or spending on that service. For example, if the specific focus of the APM is to reduce the cost of hip replacement surgery, but post-acute care providers happen to raise their prices for rehabilitation services at the same time for reasons unrelated to the changes in hip replacement surgery, an episode spending measure might show no savings or an increase in spending even if the surgeons were successful in reducing the cost of the surgery itself.

Challenges in Separating the Signal from the Noise in Composite Measures

A serious problem with using any kind of composite measure – total cost of care, episode spending, condition-based spending, etc. – is that random variation in the individual components of the measure makes it more difficult to determine what impact the APM is having in the specific areas where it is expected to produce savings. Unless the changes in those aspects of spending are very large, they may not have a significant impact on the overall composite measure, making the APM or a provider participating in the APM appear unsuccessful even when they have, in fact, successfully achieved their specific goals.

In general, the higher the percentage of the composite measure that is composed of unrelated services the APM participant cannot control, the less likely it will be that changes in the composite measure will accurately reflect changes in the types of services the APM was designed to affect or the actions of the APM participant. If there are no changes in the unrelated services, they will dilute the effect of the changes in related services, as illustrated in the example. If there are changes in the unrelated services, they will likely not be due to anything the APM participant did or could have done, so they will either inappropriately reduce the estimated effect of the APM participant’s actions, or they will inappropriately increase the estimated effect. For example, a study of Medicaid Accountable Care Organizations in

EXAMPLE OF PROBLEMS CAUSED BY USING TOTAL COST OF CARE MEASURES

Assume that an APM is designed to reduce the rate of avoidable hospitalizations for chronic disease patients. The tables show a simplified example in which the average annual spending on hospitalizations for the patients is \$3,000, the average spending on other services for the patients is \$6,000, and the coefficient of variation for both hospital spending and spending on other services is 15% (i.e., the standard deviation of spending is 15% of the average). A 25% reduction in the rate of hospitalizations would produce annual savings of \$750 per patient, and that would be statistically significant using standard significance testing methodologies even with only 5 patients. There is no change at all in other spending for the patients. However, the 25% reduction in hospitalizations would only represent an 8.3% reduction in total spending. If one were to only evaluate changes in spending on the total cost of care, the change would not be statistically significant, and an actuary or evaluator might conclude that “no savings” had been achieved because of that.

	Spending on Hospitalizations			
	Current	Under APM	Change	% Change
Average Spending	\$3,000	\$2,250	\$750	25%
Std. Dev. Of Spending	\$445	\$334	\$111	
Coefficient of Variation	15%	15%	15%	
T Statistic for Change			3.0	
Significant at 5% Level?			Yes	

	Other Spending			
	Current	Under APM	Change	% Change
Average Spending	\$6,000	\$6,000	\$0	0%
Std. Dev. Of Spending	\$891	\$891	\$0	
Coefficient of Variation	15%	15%		

	Total Spending			
	Current	Under APM	Change	% Change
Average Spending	\$9,000	\$8,250	\$750	8.3%
Std. Dev. Of Spending	\$1,336	\$1,225	\$111	
Coefficient of Variation	15%	15%	15%	
T Statistic for Change			0.93	
Significant at 5% Level?			No	

New Jersey found that the inclusion of uncontrollable components of spending in a shared savings calculation created significant biases and errors in estimates of savings compared to the true savings that the ACO produced.¹¹⁶

Striking a Balance Between Service-Specific Measures and Composite Measures

For most APMs, the best approach will be to use a combination of both service-specific measures and composite measures based on the types of impacts on spending the APM could have. Two or three separate groups of measures or composites could be defined as follows:

1. **Potentially avoidable spending**, i.e., one or more service-specific measures for aspects of spending where the APM is intended to achieve savings. For each of these measures, specific goals for savings would be defined.

Example: The Health Care Incentives Improvement Institute developed definitions of “potentially avoidable complications” associated with particular procedures and hospitalizations and with chronic disease management.¹¹⁷ Several measures based on these definitions of potentially avoidable complica-

tions have been endorsed by the National Quality Forum.¹¹⁸

Example: 3M Information Systems has developed software to identify a series of “Potentially Preventable Events,” including “Potentially Preventable (Initial) Hospital Admissions,” “Potentially Preventable Emergency Department Visits,” “Potentially Preventable Complications,” and “Potentially Preventable Readmissions.”¹¹⁹ Some or all of these classification systems are being used by a number of states and health plans as part of accountability and payment systems.

2. **Related spending**, i.e., service-specific measures, or a single composite measure, focused on specific types of services and spending where increases caused by the APM are possible but undesirable. The goal would be no increase in utilization or spending on these measures of related spending (or an increase smaller than the savings on targeted spending).

Example: In the CMS Bundled Payments for Care Improvement - Advanced APM, providers participating in the APM are responsible for controlling spending during an episode that includes an inpa-

tient hospital stay and a 90-day period of time after discharge from the hospital. In addition, spending during days 91-120 after the hospital stay is calculated and compared to the expected level of spending during this period; if the actual spending during the 91-120 day period exceeds the spending expected during that period by an amount that is significant at the 99.5% level of confidence, the provider is held responsible for that difference.¹²⁰

3. **Unrelated spending.** If there is concern that utilization or spending could increase in other, unidentified areas, an additional broad composite measure of spending could be defined by taking an episode spending measure or total cost of care measure and subtracting the aspects of utilization or spending defined in the first two groups. No specific goal would be set for this “everything else” measure, since it would presumably be affected primarily by factors unrelated to the APM; instead, it could be monitored for statistically significant changes or changes that were significantly different than changes in the same measure for patients who were not part of the APM.

c. Measuring Utilization or Resource Use Instead of Spending

Separating the Effects of Utilization and Price

Spending on a service is a function of both the utilization of the service and the payment made for the service, so savings can be achieved either by reducing utilization, reducing the price (or using a lower-priced version of the service), or both. However, the healthcare provider who

is delivering or ordering the service may have far more control over utilization than price. For example, an APM may enable a physician to successfully reduce the rate of avoidable hospitalizations for her patients, but if the hospital increases its prices, there may be no reduction in spending. In the Medicare program, hospitals, physicians, and other providers cannot simply increase the prices of their services at will, but outside of Medicare, providers’ ability to control the price of a service will be affected by the number of providers or suppliers who are competing to deliver that service. In the case of sole-source drugs and medical devices, prices can increase unpredictably even in Medicare.

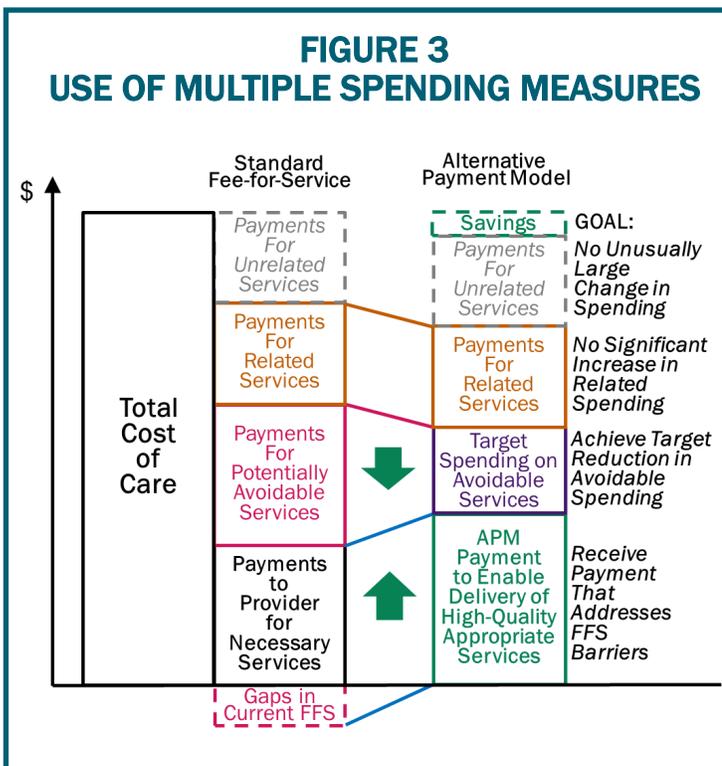
It is unreasonable for an APM to hold a provider accountable for changes in spending on a service if the provider cannot control one of the major factors affecting spending, namely, the price of the service. To address this, the APM could separately measure changes in utilization and changes in price for a service so that:

- the APM participant can be held accountable for reducing or maintaining utilization without penalizing them for price-driven changes in spending they cannot control. For example, if the hospital in the community acquires all of the independent specialist practices, the price at which a particular specialty service can be obtained may increase through no fault of the primary care physician who is ordering that service, but the PCP can still reduce unnecessary referrals to those specialists.
- The APM participant can be held accountable for spending in circumstances where it is reasonable to expect they can control both utilization and price. For example, if there are multiple laboratories in the community that charge different prices for a test, a physician could reduce spending either by reducing unnecessary utilization of the test or by using lower-priced laboratories.

Measuring Resource Use

Creating separate measures of utilization and price is problematic when there are alternative ways of delivering a service that have different prices. If a provider begins ordering a lower-priced version of a service instead of a higher-priced version, the total utilization of the service may not change (i.e., there will be offsetting changes in use of the two different versions of the service), but spending will decrease because of the lower price. A utilization measure would not accurately reflect this impact, but a measure of spending would also be problematic because increases in the prices of either of the alternative services would cause spending to increase through no fault of the provider who is ordering the services.

To address this, an artificial measure of “resource use” could be used instead of either utilization or spending. A “resource value” could be assigned to each service based on the relative cost of the service compared to other services rather than on the actual price charged for the service. The resource use for a service would then be the utilization times the resource value, rather than the utilization times the price. If it costs more to deliver service A than Service B, then service A would



be assigned a higher resource value than service B, and a shift in utilization from service A to service B would result in “savings” on the resource use measure. If prices are proportional to costs, then savings on the resource use measure would translate into savings on spending, but if the prices of one or more services are increased for reasons unrelated to their costs, the resource measure would not change even though spending would increase.

CMS uses a variation of this approach in the accountability measures it uses in APMs and in its pay-for-performance programs. CMS pays different amounts for the same service in different geographic regions because of differences in the cost of living, and it pays some providers more than others for policy reasons (e.g., it pays teaching hospitals more in order to support medical education costs). In order to separate differences in spending caused by these intentional pricing differences from differences in spending due to the use of different types of services, CMS uses a “standardized spending” measure instead of actual Medicare spending. It calculates standardized spending by assigning an artificial payment amount (the standardized payment amount) to each service rather than the actual amount that was paid for that service, multiplying the number of services used by those standardized payment amounts, and then summing the total.¹²¹

This same approach could be used to create resource use measures for payers other than Medicare, i.e., the standardized payment used by Medicare for a particular service could also be used as a measure of the relative level of resources used for that service compared to other services. An alternative is to use the relative values or weights assigned to most services in the Medicare payment system. For example, every service on the Medicare Physician Fee Schedule has a “relative value unit” (RVU) assigned to it, and every type of hospital admission has a Diagnosis Related Group (DRG) “weight” assigned to it. These RVUs and weights could be used as resource values even for services delivered to commercially-insured patients.

Example: HealthPartners, a health plan/provider in Minnesota, developed the Total Care Relative Resource Value methodology, which provides a way to measure differences in the relative levels of resources used for services and to separate the effects of pricing policies from decisions about the number and types of services used.¹²²

d. Adjusting Measures for Differences in Patient Needs

Different patients need different types and amounts of services. As discussed in Section VI.A, it will often be important to stratify the amounts of payment under an APM based on patient characteristics in order to ensure that the payments are adequate to meet the needs of individual patients. Similarly, differences in patient characteristics can also affect the measures of utilization or spending in an APM. It will be important to distinguish whether changes in a utilization or spending measure have resulted from changes in the characteris-

tics of the patients receiving services or from the impacts of changes in care delivery under the APM.

Example: Suppose that the goal of the APM is to reduce hospital readmissions, and patients who have a particular characteristic (e.g., a serious chronic disease) have historically been readmitted to the hospital 20% of the time, whereas patients without that characteristic have historically been readmitted only 16% of the time. If the APM is implemented and the overall readmission rate is 17%, it will be important to know what proportion of the patients had the characteristic which affected readmissions:

- *If all of the patients had the characteristic that is associated with higher readmission rates, then the APM has successfully reduced the readmission rate by 15% ($15\% = (17\% - 20\%) / 20\%$)*
- *If none of the patients had that characteristic, the readmission rate actually increased by 6% ($6\% = (17\% - 16\%) / 16\%$).*

Example: Suppose that the goal of an APM is to reduce overuse of knee surgery for patients with knee pain. Patients with severe pain due to complete loss of knee cartilage will likely require knee surgery in most cases, whereas patients who still have some cartilage will be more likely to benefit from physical therapy. If the APM is implemented and the rate of knee surgery increases rather than decreases, it will be essential to know if there was an increase in the percentage of patients for whom surgery is the only effective treatment.

Many APMs attempt to “risk adjust” a spending or utilization measure in an effort to separate the effects of differences in patient characteristics from the effect of changes in care delivery. In the readmission example, a “risk score” of 1.25 could be assigned to patients who have the characteristic that affects readmission, and other patients would be assigned a risk score of 1.00, reflecting the fact that the patients with the characteristic have traditionally been 25% more likely to be readmitted. The risk-adjusted readmission rate for all patients would then be determined by multiplying the actual readmission rates for each group by their risk scores and summing the products. For example, if half of the patients had the characteristic, then the baseline risk-adjusted readmission rate would be 16% ($(20\% / 1.25 + 16\% / 1.00) / 2$). After the APM is implemented, the number of readmissions would be adjusted by these risk scores to determine if there is a decrease in the risk-adjusted readmission rate.

Since there are multiple characteristics of patients that can affect their need for healthcare services, the risk score would need to be based on a methodology for combining and weighting the different characteristics to estimate what difference in utilization or spending would be expected based on the combination of characteristics each patient has. This is typically done using a linear regression model of some type, but this has the same weaknesses that were described in Section VI.A with respect to risk-adjustment of payment amounts.

In many cases, it will be simpler and better to *stratify* patients into categories based on characteristics that affect their need for services and to calculate a utilization/spending measure separately for each category. Patient categories may already have been defined in the APM for the purposes of stratifying payment amounts, as described in Section VI.A. However, because these categories were designed to align payments with the *provider's cost* of delivering the desired services, they may not adequately differentiate between patients based on characteristics that significantly affect *the payer's spending* on avoidable services. If this is so, then additional categories or subcategories of patients could be defined using those additional characteristics. The payment amounts might be the same in two of the categories, but the standard of performance on utilization or spending would be different.

Example: Patients with two different genetic variations of cancer may need to be treated with two different drugs that have very different costs, but there may be no difference in the amount of time that an oncology practice would spend in determining the diagnosis (including which genetic variation the patient has), infusing the drugs, or managing the patients' care. There would be no need to stratify patients based on the genetic variation for the purposes of paying for the oncology practice's time, but if the practice is going to be held accountable for spending on drugs, stratification on that characteristic would be necessary in order to avoid penalizing the practice for treating the patients who need the higher-cost drug.

e. Measuring Appropriateness Rather Than Utilization or Spending

The growing interest in “precision medicine” is based on a recognition that there may be a large number of factors that justify an individual patient receiving significantly more (or less) of some types of treatment than other patients. The more such factors there are, the fewer patients that have any individual factor, and the more complex the interactions among the different factors, the more difficult it will be to use either traditional risk adjustment systems or stratification systems to ensure that a provider participating in an APM is not inappropriately rewarded or penalized based on how many of the provider's patients have the specific factors affecting the need for services.

If there are generally accepted standards for determining if the service is appropriate based on each of these many factors, then the best measure may be the rate of adherence to the appropriateness standards.

In many cases, appropriateness standards can be defined for some situations but not others. There may be strong evidence that a service is or is not effective for patients with some combinations of characteristics, but only limited evidence for other types of patients. The differing levels of uncertainty are often reflected in standards by including a “may be appropriate” category in addition to “usually appropriate” and “rarely appropriate” categories. Performance can then be measured separately for each of these categories, i.e., the percent-

age of patients who received the service when it was classified as “rarely appropriate” vs. the percentage of patients who received it when it was classified as “may be appropriate.”¹²³

f. Defining the Timeframe for the Measures

In addition to defining the *types* of services for which utilization, spending, resource use, or appropriateness will be measured, a decision must also be made about the *timeframe* in which those services must occur in order to be included in the measure.

Advantages of Measuring Over Longer Timeframes

In most cases where the goal of the APM is to reduce avoidable use of a planned service or to use a lower-cost service, the expected savings will occur immediately or within a short period of time after the delivery of the services supported by the APM. For example, if the APM is designed to enable a provider to use a lower-cost alternative to an existing service, then the use of the lower-cost alternative presumably generates immediate savings by avoiding the use of the higher-cost service.

However, if a provider has discretion not only about *whether* to use or order a service but *when* to do so, there is the possibility the provider could delay using a service rather than avoiding it altogether. For example, if a physician participating in an APM begins delaying ordering tests or procedures for patients, spending may decrease during the initial measurement period, making it appear that savings have been generated, even though spending would then increase again in the next measurement period.

Example: In the CMS Oncology Care Model, spending is measured in six-month episodes following the date that a patient receives chemotherapy to treat cancer. If the patient receives chemotherapy treatments for more than six months, a new six-month episode is triggered, and the total spending for the patient during their chemotherapy treatment is divided between the two episodes.¹²⁴ This means that if a patient would ordinarily have completed their treatment in six months, but the oncology practice delays one or more treatments to the seventh month, the total spending on the patient would be divided into two episodes rather than one, making the average spending per episode appear lower than it would otherwise.

When the goal of the APM is to reduce *unplanned* services, it is likely that some or all of the changes in spending associated with an APM will occur days, weeks, months, or even years after the desired service is delivered. For example, prevention and early treatment of diabetic foot ulcers can avoid the need for amputations, but the improvements in diabetes care need to begin long before an amputation would be necessary – in many cases, years before – so the savings from avoided amputations will occur much later than any increase in spending on improved services. In some

cases, savings will accumulate over time; for example, improved care management services delivered during a particular month may help a patient with chronic disease avoid a hospitalization during the same month, but it may also help them avoid hospitalizations in future months as well.

To address these issues, a utilization or spending measure in an APM can be defined in terms of a specific period of time after the services supported by the APM are delivered. For example, a measure of utilization or spending on hospital readmissions could be defined based on readmissions that occur within 60 or 90 days following hospital discharge.

Disadvantages of Measuring Over Longer Timeframes

Although the considerations above make it appear desirable to use the longest time period possible for measuring spending, the greater the lag in time between when a service is delivered and when utilization or spending is measured, the greater the possibility that changes in utilization or spending have occurred for reasons unrelated to the original service. For example, the percentage of patients who are readmitted to the hospital is higher if readmissions are measured over a longer period of time following discharge (e.g., 90 days instead of 10 days), but a smaller percentage of the readmissions are directly related to the original reason for hospitalization.¹²⁵

In addition, there are practical problems involved in implementing APMs using measures of spending defined over a long period of time. If there is a large delay between the time that the service supported by the APM is delivered and the time a determination can be made as to whether the intended savings have been achieved, there will either need to be a delay in paying for the desired service under the APM until the desired performance has been confirmed, or the payer or patient will need to pay for delivery of the desired service immediately and then seek recoupment later if it is determined that the spending target was not met. Both approaches can be problematic for small providers, who may not have sufficient financial reserves to deal with delays in payments and who may not have the data or

actuarial skills to build reserves for potential recoupments.

Using Multiple Measures With Different Timeframes

One way to balance the advantages and disadvantages of longer vs. shorter timeframes is to use multiple measures with different timeframes. For example, if the goal of the APM is to reduce hospital readmissions, the performance component could use two measures of readmissions: (1) readmissions that occur within 30 days and (2) readmissions that occur within 90 days. Different performance standards and penalties could be assigned to each measure in order to reflect differences in a provider's ability to reduce readmissions earlier and later and also to reduce cash flow problems for the provider.

Example: In the Bundled Payments for Care Improvement-Advanced APM, CMS has defined two different time periods for measuring spending related to a hospital admission. The first time period is the 90 days immediately after discharge, and the second time period is the next 30 days after that, i.e., days 91-120.¹²⁶

**TABLE 8
ALTERNATIVE WAYS OF MEASURING UTILIZATION AND SPENDING**

Measure	Strengths	Weaknesses	Examples
Spending for a Specific Service the APM is Intended to Reduce	Focuses on exact services APM is intended to reduce	Service may be delivered to patients not included in the APM; provider may be able to control utilization but not price of service; spending may be higher if provider has a higher-than-average number of patients with characteristics requiring delivery of the service	Spending per patient on ED visits
Spending for a Service in Specific Circumstances	Focuses on circumstances in which APM is intended to reduce spending	Requires ability to define and identify existence of circumstances	Spending per patient on ED visits related to the chronic diseases targeted by the APM
Spending for a Group of Services APM is Intended to Reduce	Adjusts for offsetting changes across all services APM is intended to reduce	Use of other services may change in unexpected ways	Spending per patient on ED visits and hospitalizations related to the chronic diseases targeted by the APM
Spending for All Services the Patient Receives During a Period of Time	Captures all changes in services	Net change may be due to unrelated factors the provider cannot control; random variation in unrelated services may mask targeted changes	"Total cost of care" per patient "Episode spending" per patient
Multiple Measures of Spending	Allows different Targets to be used for each measure	Requires clear definitions for what types of services are included in each measure	Measures of three subsets of spending: - Avoidable Spending - Related Spending - Unrelated Spending
Utilization of a Specific Service	Focuses on change in use of service, not changes in price	Does not encourage use of alternative services that have lower prices	Number of ED visits per patient that are related to the chronic diseases targeted by APM
Resources Used for a Group of Services	Focuses on extent to which higher-cost or lower-cost services are used but not on differences in price relative to cost	Does not encourage use of lower-cost providers	# of ED visits multiplied by a standardized cost per ED visit + # of hospital admissions times standardized cost per admission) divided by # of patients
Risk-Stratified Spending on a Service	Controls for patient characteristics that require more services or more expensive services	Reliable data on important patient characteristics may not be available	Spending per patient on ED visits for (1) high-risk patients, (2) medium-risk patients, and (3) low-risk patients
Utilization/Spending on Services Not Meeting Appropriateness Criteria	Focuses on the services that can be eliminated without harming patients	Reliable data on patient characteristics needed to assess appropriateness may not be available	Spending per patient on ED visits for (1) low-mortality, non-urgent symptoms and (2) urgent/high-risk symptoms
Spending During Different Timeframes	Separates short term and long-term spending and savings	If boundaries of time periods are arbitrary, spending could shift between time periods	Spending per patient on ED visits during (1) the first 30 days and (2) the next 60 days

2. Setting the Performance Targets for Utilization and Spending

As discussed in Section IV.C, the business case for the Alternative Payment Model will be based on an expectation that the savings to be achieved from reducing one or more types of services or spending will exceed any higher costs and higher payments required to support an improved approach to care delivery. The business case could be undermined if the APM fails to achieve the expected savings or unintentionally results in increased spending in other areas.

Consequently, a Target for each of the measures of spending or utilization identified in the previous section will need to be established based on what is needed to assure that the business case for the APM is being fulfilled.

a. Patient-Level Targets vs. Population-Level Targets

i. Targets for Planned Services

Some APMs that produce savings primarily through changes in planned services can be expected to achieve net savings for every participating patient (or every patient who has specific characteristics), either because those patients will no longer receive an unnecessary service or because any additional spending on services for those patients will be offset by savings from lower utilization or spending on one or more other services for the same patients. For example, if an APM pays for a new service A that is designed to substitute for a more expensive service B, then spending will be lower on each patient who receives service A instead of service B. Similarly, if the APM enables a similar service to be delivered in a less expensive setting or in a less expensive way, then spending will be lower on each patient who receives the service in the new way. In some cases, the substituted service may not be any less expensive, but if it achieves better outcomes for the patient, it could qualify as an APM under MACRA.

In these cases, a *Patient-Level Target* is most appropriate, i.e., the Target is the amount that is expected to be spent on each individual patient. For example, if the goal of the APM is to avoid use of a particular service, the patient-level Target could be defined as zero utilization or spending on that service for each patient participating in the APM or for a subset of patients with particular characteristics. If there are multiple subsets of patients with different types of services to be avoided or substituted, then either:

- separate Targets could be defined for each category of patients defining the specific utilization/spending changes expected for that category, or
- the Target could be defined as adherence to evidence-based clinical guidelines or “pathways” that specify which services are appropriate for each patient.

Other APMs will be designed to enable planned services to be delivered in different ways, but the amount of spending will vary from patient to patient, and the goal will be to reduce the average spending per patient on planned services, not necessarily to reduce the spending

on each individual patient. In these cases, a *Population-Level Target* will be needed, i.e., the average level of spending or rate of utilization that should be achieved overall for the patients participating in the APM. The Population-Level Target tells the patient what their “expected spending” would be in probabilistic terms, but for any individual patient, the actual spending may be higher or lower than that amount.

ii. Targets for Unplanned Services

In APMs that are designed to reduce unplanned services, if there is an expectation that the unplanned service can be avoided in all, or almost all, cases, then a Patient-Level Target can be used. For example, if there is evidence that a particular type of hospital-acquired infection can be prevented in close to 100% of cases by delivering a set of services supported by the APM, then a Patient-Level Target for spending on that infection could be set to zero.

For many types of unplanned services, the expectation will be that the rate of occurrence will be lower but the services will not be eliminated entirely. For example, patients who have many types of chronic disease, such as heart failure and COPD, are at risk for hospitalizations due to exacerbations of their disease. The services supported by the APM may be expected to reduce the rate of such hospitalizations, but not to eliminate them entirely, and there may be no way to reliably determine in advance which patients will and will not be hospitalized.

In general, a *Population-Level Target* will be needed in these situations, i.e., the Target would be equal to the average level of spending on the unplanned services or the rate of utilization of such services that should be achieved overall for the patients participating in the APM. If the APM is designed to reduce hospitalizations, the Population-Level Target would be defined as the average amount to be spent on hospitalizations for all of the patients receiving services supported by the APM. As with planned services, the Population-Level Target tells the patient what their “expected spending” on unplanned services would be (i.e., the probability of a hospitalization times the payment for an individual hospital admission), but for any individual patient, the actual spending will be higher than that Target (if a hospitalization actually occurs) or lower than the Target (if no hospitalization occurs), it will never be exactly equal to the Target.

iii. Advantages of Patient-Level Targets

An APM that uses Patient-Level Targets is preferable from the perspective of a patient who is paying for their own care, because they will know how much they will have to spend, whereas under a Population-Level Target, an individual patient could have to spend much more than the Population-Level Target and/or spend more than they would have spent in the absence of the APM, as long as sufficient savings are achieved for other patients so that the overall average is lower. From a payer’s perspective, Patient-Level Targets are not inferior to Population-Level Targets, because if spending is the

same or lower for each *individual* patient in a group insured by a payer, then spending will also be lower for the group as a whole.

In some cases, Patient-Level Targets can be used instead of Population-Level Targets if the patients can be stratified into groups that have more predictable levels of spending, as discussed in Section VI.B.1.d. For example, if it is known that certain patients will require more planned services than others, or they are at higher risk of unplanned services, a separate Target could be established for them. This is preferable to “risk adjusting” a Population-Level Target as many APMs do, because the risk-adjusted Population-Level Target does not tell either the patient or the provider what the spending goal is for an individual patient.

iv. Targets for Composite Spending Measures

Composite measures such as “episode spending” and “total cost of care” measures typically include spending on both planned and unplanned services. If Patient-Level Targets can be established for the individual types of services, then a Patient-Level Target can also be defined for the composite, but if a Population-Level Target would be needed for any individual type of service that is included in the composite, then it is likely that a Population-Level Target would have to be used for the composite itself.

Some APMs have described Population-Level Targets for composite measures as though they are Patient-Level Targets when they are not. For example, in the Bundled Payments for Care Improvement (BPCI) – Advanced APM, CMS defines a “Target Price” for each episode based on the average spending on patients with similar episodes in the past. However, the Target Price is only used to calculate the “Target Spending” level for a *group* of patients, and the success or failure of a provider participating in the APM is determined based on whether *total* episode spending for the *group* is more or less than the Target Spending amount. There is no penalty if spending for an *individual* patient exceeds the Target Price as long as there are enough other patients for whom spending is below the Target Price, and no provider is eligible to participate in the APM for a particular type of procedure unless the hospital has at least 40 patients receiving that procedure.¹²⁷ This is a Population-Level Target, not a Patient-Level Target.

Even if a Patient-Level Target cannot be defined for every service in a composite measure, that does not mean the only option is to use a Population-Level Target. If patients could be stratified into subgroups where spending on the composite measure in each subgroup is predictable, then a separate Patient-Level Target could be established for each subgroup. Alternatively, the subset of services that are less predictable could be separated from the composite, so that a Patient-Level Target can be established for the predictable services and a Population-Level Target would then be established for the less predictable services. This is a variation on the approach described in Section VI.B.1.b for using both service-specific and composite measures with different Targets.

In addition, if a true Patient-Level Target can be defined for spending on a group of services, i.e., the amount that should be spent for each individual patient, the provider

could be paid a bundled payment with that amount as part of Component #1 instead.

b. Alternative Ways of Setting Population-Level Targets

There are several ways in which Population-Level Targets for spending (or utilization or resource use) can be defined:

- Benchmark-Based Targets;
- Evidence-Based Targets; or
- Competitive Targets.

As will be discussed below, Benchmark-Based Targets depend on having information about utilization and spending levels for similar patients who are receiving services from providers who are not participating in the APM. The more widely the APM is used and the longer it is used, the more difficult it will be to develop good Benchmarks, and the more likely it will be that Evidence-Based Targets, Competitive Targets, or other methods of setting Targets will be needed.

i. Benchmark-Based Population-Level Targets

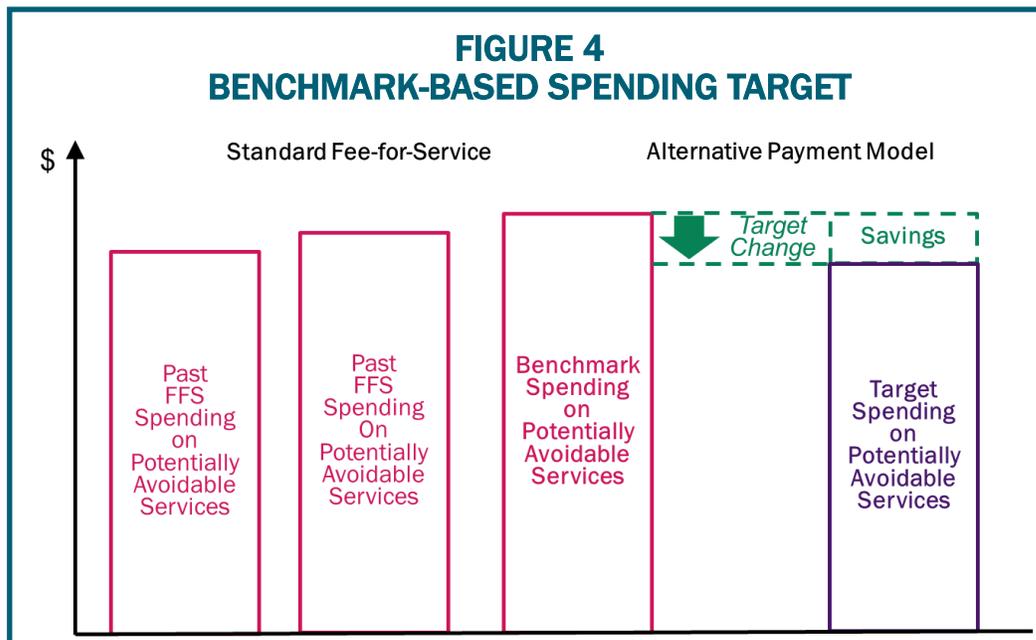
Because spending under an APM is required to be equal to or lower than it would have been in the absence of the APM, a common approach is to define the Target relative to what utilization or spending is currently or is expected to be in the absence of the APM. This requires specifying two separate components:

1. a population-level *Benchmark* that defines what level of spending/utilization for the patients receiving services supported by the APM is viewed as reflecting “no impact of the APM”; and
2. a *Target Change*, i.e., the minimum or maximum amount by which actual spending or utilization under the APM should differ from the Benchmark.

Example: *If the APM is paying more to a physician practice to deliver home care services with a goal of reducing avoidable hospitalizations, then the Benchmark would be the rate of avoidable hospitalizations or the amount of spending on such hospitalizations that is expected in the absence of the APM, and the Target Change would be the reduction in avoidable hospitalizations needed to generate sufficient savings to cover the payments for the home care services.*

The term “benchmark” has generally been used differently when applied to utilization/spending measures than to quality measures. For quality measures, the term “benchmark” has typically been used to refer to a *desirable* level of performance (e.g., the best performance that has been achieved to date), in contrast to utilization/spending measures, where “benchmark” has been used to refer to the *status quo* performance level. Both of these uses are consistent with the dictionary definition of “benchmark,” which is “a standard or point of reference against which things may be compared or assessed.” Since APMs will generally include both utilization/spending and quality measures, it will be im-

**FIGURE 4
BENCHMARK-BASED SPENDING TARGET**



portant to clarify what is meant by a “benchmark” in each case.

For consistency, this report will use the term “Target” to describe the “desirable level” for both utilization/spending and quality measures. The Target is the key parameter in defining the accountability component, and a benchmark is merely one way of determining the Target.

ii. Evidence-Based Targets

If there is evidence indicating that a specific level of utilization or spending can be achieved that is lower than the level currently being achieved by most providers, then that level of utilization and spending could be set as the Target, thereby avoiding the need to define Benchmarks and Target Changes. For example, if the accountability measure used in the APM is spending for treatment of complications such as infections, and if evidence has shown that a specific low rate of infections and other complications can be consistently achieved for patients being treated for a particular condition, then the level of services or spending needed to treat that rate of complications could be set as the Target for the APM.

The APM itself may facilitate the development and use of Evidence-Based Targets. If an APM is initially implemented using a Benchmark-Based Target and if a wide range of APM participants show that they can consistently achieve levels of utilization or spending that are lower than the Benchmark, the actual performance levels on utilization and spending could be used as an Evidence-Based Target.

iii. Competitive Targets

In situations in which there are multiple providers offering services under an APM, the Target could be set through a competitive process, instead of by defining Benchmarks and Target Changes. Each provider would define the Target level it was willing to be held accountable for achieving (either for individual patients or on aver-

age for a group of patients), and payers or patients could then choose a provider based in part on a comparison of the Targets defined by the different providers. The provider who set the lowest Target on a particular aspect of utilization or spending would not necessarily be the preferred choice, since that would depend on how much the provider was charging or being paid for the services being delivered, the level of quality the provider had committed to achieve, the Targets for other aspects of utilization and spending, the provider’s past record of achieving the Targets, etc.

c. Alternatives for Defining Population-Level Benchmarks

If a Benchmark-Based Target is going to be utilized, three basic methods can be used to define the Benchmark:

- A Prior Performance Benchmark, defined either for the same patients or the provider’s patients in the past;
- A Comparison Group Benchmark; or
- A Counterfactual Benchmark

i. Prior Performance Benchmark

In this method, the Benchmark is based on the actual level of spending or utilization during a previous period of time, e.g., the prior year or an average of several prior years. Two different variations of this approach can be used, depending on the circumstances:

- **Prior Performance for the Same Patients.** If the same patients will be receiving services under the APM for the same conditions for which they had received services in the past (e.g., if the APM is focused on wellness, preventive care, chronic disease management, or other services that recur over time), then the Benchmark for the patients could be based on actual spending or utilization in the past for those same patients. For example, if the APM is designed to reduce spending on hospitalizations for patients with a chronic disease, then the prior year’s spending on hospitali-

zations for the patients could be used as the Benchmark for the current year's spending on hospitalizations for those same patients.

- **Prior Performance for Similar Patients.** If the APM is focused on acute conditions, the patients receiving services under the APM will not have received similar services before (or at least not under similar circumstances). Here, the Benchmark could be defined as the provider's level of utilization or spending during the prior year for a similar group of patients in similar circumstances. For example, if the APM is designed to reduce spending on patients who have a heart attack, the Benchmark could be based on how much the provider spent treating similar patients who had a heart attack in the past.

ii. Comparison Group Benchmark

In this method, a group of patients is identified who are not participating in the APM, but who are similar to those who are in the APM, and the Benchmark is based on their actual spending or utilization during the same period of time as that in which the APM services are delivered. For example, if the APM is designed to reduce spending on hospitalizations for chronic disease patients, the Benchmark would be the amount of spending for similar patients with the same disease who are not part of the APM.

iii. Counterfactual Benchmark

A third method is to attempt to estimate what the spending or utilization in the current year would be for the specific patients who are receiving services supported by the APM. A Counterfactual Benchmark will often be based in part on prior utilization/spending for the participating patients or utilization/spending of a comparison group, but other adjustments will be included to address expected levels of inflation in costs, new treatment technologies or changes in evidence about what treatments are most effective, etc. For example, if the APM is intended to reduce spending on hospitalizations for patients with a chronic disease, a model that predicts the probability of hospitalization for individual patients with that disease could be used to establish a Benchmark rate of utilization for the group of patients, and if a high level of inflation in the national economy is pushing costs higher, the Benchmark could be increased by the projected rate of inflation.

iv. Combinations of Benchmark Approaches

Many current APMs use combinations of these approaches. For example, in the CMS Oncology Care Model, the Benchmark for each participating oncology practice is determined through a complex process that includes:

- Estimating the amount that would have been spent for each cancer patient using a regression analysis calibrated on past spending (i.e., basing the Benchmark in part on a counterfactual);
- Adjusting the spending estimate using an "experience adjuster" based on the individual practice's past spending relative to the average of other practices (i.e., basing the Benchmark in part on prior performance); and

- Increasing the spending estimate for the cost of "novel therapies" if the individual oncology practice uses such therapies at a higher rate than other oncology practices (i.e., basing the Benchmark in part on a comparison group).¹²⁸

d. Alternatives for Defining Target Changes

Since the Benchmark for a measure is intended to represent the level of spending/utilization that reflects "no impact" of the APM on spending or utilization for that particular service or combination of services, the Target Change will define the magnitude of the desired impact of the APM. There are four different approaches that could be used to define the Target Change:

- Minimum/maximum change needed for success;
- Change achieved by a comparison group;
- Statistically significant change; or
- Desired level of change.

i. Minimum/Maximum Change Needed for Success

If the APM is intended to reduce utilization or spending on a particular measure compared to the Benchmark, the Target Change could be set at a level that is at least sufficient to achieve adequate savings to offset any expected increases in spending on desirable services and/or sufficient to justify the payer's participation in the APM.

For a measure where no savings is expected and the goal is to avoid an increase, the Target Change could be defined as either:

- zero; or
- an increase that would be less than the net savings expected from the planned services and the intended reductions in avoidable services under the APM.

Example: If the APM pays a physician practice \$50 per patient per month to support improved care management services for heart failure patients with a goal of reducing avoidable hospitalizations, the Target Change for reduction of avoidable hospitalizations could be set at an amount equal to or greater than \$50 per patient per month. In order to assure that a reduction in avoidable hospitalizations is not being achieved by greater use of home health services, the Target Change in home health services could be defined as \$0 per patient per month.

If the increased spending on services supported by the APM will differ for different patients, then the Target Change may also need to vary in the same way, or alternatively it could be based explicitly on the amount by which spending differs. For example, if higher care management payments are paid for patients with more complex conditions and higher rates of hospitalization, then the Target Change (i.e., the desired savings in hospitalizations) could be larger for those same patients.

ii. Change Achieved by a Comparison Group

Although it would be desirable to achieve greater savings than the minimum amount needed for success, it may be unclear exactly how much savings is possible. There is frequently considerable uncertainty regarding whether unplanned care will occur and the extent to which changes in planned care will be able to affect it. One approach to addressing these uncertainties is to define the Target Change based on what other participants in the APM have achieved, or what participants in other initiatives have achieved. This could either be the average performance of all participants in the comparison group or the performance of a higher-performing subset. For example, the Achievable Benchmarks for Care (ABC™) methodology determines the best performance level that has been achieved for at least 10% of patients by a group of providers.¹²⁹

iii. Statistically Significant Change

There is a considerable amount of patient-to-patient variation in utilization and spending on services. Some of this variation is controllable by the APM participant, some of it is not controllable but predictable (because of differences in patient characteristics or differences in unit costs in different communities that lead to systematic differences in spending), and some variation is neither predictable nor controllable and is essentially random from the perspective of both payers and providers. If an APM participant's spending is lower than expected because of random variation rather than because of the effect of the APM, it would be inappropriate to say that the APM participant was successful in achieving savings, and if the participant's spending is higher than expected due solely to random variation, it would be inappropriate to say that the participant was unsuccessful.

To address this, the Target Change could be defined in such a way as to provide confidence that the change was not due to random variation. If a minimum amount of savings for a particular service or group of services is needed to support the business case for the model, then the Target Change for those services could be set at a level sufficiently higher than the minimum to provide assurance that the "true" savings was at least equal to the minimum. If the goal is to avoid an increase in spending on a group of services, the maximum increase permitted may need to be greater than zero to reflect the fact that random variation could cause spending to appear to be higher than the Benchmark in any given year.

If the Target Change is defined in this way, a decision has to be made about the desired level of confidence that a change is not random. For example, if the goal of the APM is to achieve savings of at least \$100 per patient on a particular measure, the Target Change would need to be set much higher than \$100 in order to achieve a confidence level of 95% (which means there is only a 5% probability the savings are less than \$100) and the Target Change would have to be set higher for a 95% confidence level than if only 90% confidence was needed. However, the higher the confidence level required to declare that savings have been achieved, the higher the probability that an APM participant will be inappropriately deemed to have failed. (For example, even if spending decreases by \$100 per patient, it might not be deemed a

success because there *could* have been a reduction of \$100 due to random variation when the "true" savings was \$0.) The tradeoff between "Type I" and "Type II" errors in measuring performance is discussed in more detail in Section VI.B.3.

iv. Desired Level of Change

The Target Change amount could be set at a level that would achieve a specific amount of savings that is desired by the payer or the providers or to achieve a specific level of utilization or spending that is believed to be desirable and achievable. This amount would presumably be at least as much as the minimum change needed to produce some savings (i.e., the first approach), but it could also be much larger than that if more savings are needed, if the opportunities to achieve savings are large, or if there is evidence that much lower levels of utilization and spending can be achieved.

If the amounts of payment under the APM are stratified as described in Section VI.A (i.e., the amounts differ between the categories), then it will generally be desirable to also stratify the Target Changes in each category so they align with the differences in payment. For example, if the APM pays for a care management service that is designed to reduce avoidable hospitalizations for a patient with a chronic disease, and if patients whose disease is more severe have a higher rate of hospitalizations, a higher Target Change for avoidable hospitalizations could be established for those patients.

Examples of Target Changes in Current APMs

Current Alternative Payment Models use different approaches to setting the Target Change. For example:

- In the BPCI-Advanced APM, CMS uses the Desired Level of Change option, requiring that spending on an episode of care be at least 3% below the benchmark.¹³⁰
- In the Medicare Shared Savings Program, CMS uses the Statistically Significant Change option, requiring that an ACO achieve a Minimum Savings Rate in order to receive a shared savings payment. The Target Change (i.e., the Minimum Savings Rate) is larger for ACOs with fewer assigned patients because a larger change is needed to achieve the same level of statistical significance, but the difference in the Target Change is smaller than it would otherwise be because lower levels of confidence are also used in determining the Target Change for smaller ACOs.¹³¹
- In the Independence at Home APM, CMS uses the Statistically Significant Change option, requiring that a participating physician practice achieve a spending reduction large enough to achieve a minimum level of statistical significance. In this APM, there are two different Target Changes defined, with different payments associated with achieving them: If the spending reduction is significant at the 10% level (i.e., there is 90% confidence that the change in spending is not due to random variation), the practice receives 50% of the savings, but if the spending reduction is significant at the 5% level (i.e., there is 95% confidence that the change was not random), the practice receives 80% of the savings. Because it is harder to achieve

the same level of significance with fewer patients, practices with fewer patients have to achieve bigger reductions in spending; for example, a practice with 200 patients has to reduce spending by 11.4% to meet the 10% significance threshold and receive 50% of the savings, whereas a practice with 2,000 patients only has to reduce spending by 3.5% to receive a share of the savings.¹³²

e. Issues in Defining Targets

i. Prospective vs. Retrospective Targets

In addition to specifying *how* Targets will be determined, the APM needs to specify *when* they will be determined. A *Prospective Target* is set using a Benchmark and Target Change that are determined before the beginning of the time period in which performance is going to be evaluated. A *Retrospective Target* is set using a Benchmark and/or Target Change that is determined during or after the end of that time period.

There are advantages and disadvantages to each approach:

- Prospective Targets enable the APM participant to clearly understand the level of performance that is required before care delivery begins and to make adjustments in care delivery if actual performance is falling short of the Target. In contrast, under a Retrospective Target, it will be difficult for a provider to make adjustments in services since it won't know for sure how it's doing until after the care has already been delivered.
- Prospective Targets enable the payer to better predict its spending under the APM.
- Prospective Targets require making assumptions about trends in prices or differences from comparison groups that may not prove to be accurate. The provider of services may have more difficulty meeting the Target than expected, or the payer may not actually experience savings, simply because the Target was not set correctly. Retrospective Targets can be more realistic because they are based on actual data.
- Lags in the availability of data used to define Benchmarks mean that Prospective Targets may have to be set based on utilization and spending measured at a time when relevant aspects of care were delivered in different ways or when different types of patients were receiving care. For example, if annual measures of spending are used, the Benchmark for the current year will generally need to be based on actual spending from two years earlier.¹³³ However, these same data lags also mean that Retrospective Targets cannot be established until long after the services are actually delivered.¹³⁴

For example, in the Medicare Shared Savings Program, the Target spending level for ACOs (which CMS refers to as the "benchmark") has to be established retrospectively because the methodology uses the actual increase in national Medicare spending during the performance year to trend the ACO's historical spending forward, and the actual increase is only known after the performance year has ended.¹³⁵ In order to create a Prospective Tar-

get, a forecast of the increase in growth would need to be made, and that forecast could turn out to be too high or too low.

ii. Participant-Specific Targets vs. Common Targets

A decision must also be made as to whether the *same* Benchmark and/or Target will be used for *all* providers participating in the APM, or whether the Target will be customized for each APM participant.

For example, many current APMs have been designed as "shared savings" models, in which the amount an individual entity (e.g., a physician practice or hospital) participating in the APM is paid depends on whether spending is less than expected for *that specific entity*. This requires creating a different Target for each individual APM participant.

It is clearly easier for a payer to assure that overall savings are being achieved in an APM if each individual participant is required to achieve savings. However, this approach is problematic because it can reward providers who have had higher-than-average levels of avoidable spending in the past and who then reduce that spending under the APM, while penalizing providers who had already found ways to reduce avoidable spending prior to the APM but might need to use the APM to sustain those approaches. It also paradoxically means that individual patients and payers could be paying more for care from APM participants that have "achieved savings" than if they had received the same services from APM participants that "increased spending."¹³⁶

The participant-specific approach is required for Accountable Care Organizations (ACOs) under the Medicare Shared Savings Program because the law requires that the Benchmark for *each* ACO be based on historical spending for beneficiaries assigned to *that* ACO, and that an ACO can only receive a shared saving payment if the average spending for beneficiaries assigned to the ACO is lower than the benchmark for *that* ACO.¹³⁷ However, none of the other statutes authorizing the creation of APMs require that the payment for each individual participant in an APM should depend on whether *that specific provider* reduced spending. For example, the authorizing statute for the Center for Medicare and Medicaid Innovation requires only that the total Medicare or Medicaid spending under the APM be no greater than it would have been if the APM had not been implemented.¹³⁸

Advocates of participant-specific Targets argue that if the same Target were used for all participants, the majority of providers who would participate in the APM would be those whose performance is already equal to or better than the Target, and this would reduce the resulting savings or even lead to increased spending (because an increase in a very low level of spending might still be below the Target). However, this assumes that if the providers who already have low levels of spending didn't participate in the APM, they would continue to have low levels of spending. There is nothing in the fee-for-service system that assures that would occur, whereas participation in the APM could both *enable* the low-spending providers to keep spending at low levels

and to assure they would do so, even if they do not produce *additional* savings. Moreover, if the APM is specifically designed to pay adequately for appropriate services to patients and if a common spending Target for avoidable services is set at a level that is achievable, then participation should be attractive to providers that are not achieving the Target as well as those that are. Even if it turns out that the initial participants in the APM disproportionately consist of those that already have low levels of avoidable spending, if they find that participation in the APM allows better care to be delivered, it may help to attract additional participants.

iii. Comparability of Patients in Defining Benchmarks

A key issue in all of the options for Benchmark-Based Targets is making sure that the patients used in calculating Benchmarks are similar to the patients participating in the APM and making appropriate adjustments for any differences. Even if the Benchmark is based on prior utilization/spending for the same patients, the patients may have changed in ways that would cause utilization or spending to change in the absence of the APM. For example, if the APM is focused on patients with a chronic disease, the severity of a patient's chronic disease will generally increase over time, resulting in more hospital admissions and other problems. Also, if utilization or spending was randomly low or high during the previous period, then regression to the mean would result in an increase or decrease in spending in the current period that has nothing to do with the APM.

If the measure of utilization or spending is risk-adjusted or risk-stratified to address differences in patients, then the Benchmark will also need to be adjusted or stratified using a similar methodology, and this will help in improving the fairness of the Benchmark.¹³⁹ However, if the APM is focused on patients who have specific characteristics, and information about those characteristics is not routinely or reliably collected by providers under current payment systems, that will introduce additional errors and uncertainty into the Benchmarks. For example, if the APM is intended to reduce complications of treatment for a very specific condition (e.g., Stage IV breast cancer), the Benchmark should be based on the expected level of spending on complications for that specific condition, since spending on complications for other conditions (e.g., earlier-stage breast cancer) will likely be different. However, since standard healthcare claims data do not record all of the key characteristics of individual conditions (e.g., the stage of cancer is not recorded in ICD-10), it will be impossible to measure how much has been spent on complications of that particular condition in the past and it will be more difficult to define a comparison group in the current year.

If the gaps in data on comparison populations will result in inaccurate Benchmarks, it will be necessary to either find a way to obtain the necessary data (e.g., by paying to collect it from providers who are not participating in the APM) or by changing the way the Benchmark is defined.

iv. Diagnosis-Finding and Upcoding

A further problem is created if the measure of utilization or spending is risk-adjusted based on diagnoses or other patient characteristics that do not necessarily require treatment or for which payments for treatment can be received under the fee-for-service system without documenting the presence of the diagnosis. For example, if physicians receive higher payments under the APM for patients with more comorbidities, the physicians participating in the APM will have a financial incentive to document the presence of comorbidities and potentially to record diagnoses even if the patients do not need or want treatment for those conditions. In contrast, physicians in the fee-for-service system are paid based on the procedures they perform, not on how many comorbidities their patients have, so they will be less likely to record all health problems the patients have. This can make patients in a comparison group appear healthier than patients in the APM even if they are actually identical, which in turn will make risk-adjusted spending in the APM appear to be lower even when it is not.

Payers have used problematic approaches to address this problem, such as imposing an arbitrary cap on the amount by which the average risk score for a group of patients is allowed to increase.¹⁴⁰ This penalizes providers whose patients have legitimately become sicker. Instead, the following methods can be used to minimize the problem:

- **Use more narrowly focused spending measures.** Like many other problems described earlier, this problem can be mitigated by using a measure that is focused on the types of utilization or spending that are believed to be avoidable by actions the APM participant can take. A measure of total spending can be affected by every conceivable diagnosis, but a measure focused on a particular aspect of avoidable spending will likely only be affected significantly by a narrower range of conditions.
- **Require that the patient receive treatment for the diagnosis or receive specialized services to address a particular characteristic in order to include that diagnosis or characteristic in the risk-adjustment system.** The purpose of including a diagnosis or other characteristic in a risk-adjustment system is to capture the magnitude of the effect of that diagnosis/characteristic on utilization or spending. If the patient does not need to be treated for the diagnosis, or if the services being delivered to the patient are not different in any way from the services other patients receive, then the presence of the diagnosis will have no effect on spending. However, if the provider and patient have made an explicit decision not to treat a condition that will affect spending or outcomes (e.g., in an end-of-life situation), then that decision could be documented to justify using the condition in the risk-adjustment methodology.
- **Adjust the Benchmark when pre-existing conditions are identified.** If the APM participant has documented for the first time that a diagnosis or other characteristic exists, but the condition is not new to the patient but merely newly documented (i.e., the patient had the condition in the past, but it was not being record-

ed or treated), then any utilization or spending information regarding that patient that was used to create the Benchmark should be adjusted to reflect the presence of the condition. For example, if the Benchmark is based on the prior level of spending on the same patients, if the presence of a particular disease is viewed as a significant factor affecting spending in the APM, and if some patients who clearly have had that disease for some period of time have the diagnosis documented for the first time under the APM, the Benchmark spending amount (or the assessment of performance) should be recalculated to reflect the fact that those patients had the disease in the past.

v. Target Changes vs. Benchmark Adjustments

It is important to distinguish between the Target Change – which is the *desired* change in utilization/spending under the APM – and any adjustments made in calculating the Benchmark in order to reflect changes that would have been expected to occur *even in the absence of the APM*. For example, if there is no change in actual spending on a group of patients under the APM, it could either be because spending would have been expected to increase significantly even in the absence of the APM, and the providers achieved savings that were just large enough to offset that increase, or it could be because spending was not expected to increase and the APM participant failed to achieve any savings. If a Counterfactual or Comparison Group Benchmark were used, it would indicate whether an increase in spending would have been expected in the absence of the APM, and the Target Change would indicate how much of a change in actual spending the provider would be expected to achieve.

CMS has been criticized for underestimating the savings from ACOs in the Medicare Shared Savings Program (MSSP) because savings calculated by comparing spending with the benchmark established for ACOs is lower than what some researchers have estimated by comparing spending to a comparison group.¹⁴¹ However, the “benchmark” spending level in the MSSP is not intended to be a true counterfactual Benchmark; it is actually a Target spending level, because the methodology CMS uses to set the benchmark adjusts spending in a way that is designed to reduce the differences in spending among different ACOs over time. After calculating the per-beneficiary spending level for each ACO in previous years, CMS adds the dollar amount of the national average increase in spending per beneficiary during the performance year. This means that the benchmark will increase less in percentage terms for an ACO that had above-average spending in the past than for an ACO that had below-average spending. As a result, an ACO with above-average spending in the past will have to decrease spending more in percentage terms in order to achieve “savings,” which is defined as a difference between actual spending and the benchmark. It also means the benchmark will increase less than the increases in spending for providers in a different geographic area where utilization of services increased more than the national average.¹⁴²

vi. Revising Targets and Changing the Target Methodology Over Time

Once a Target is set, a decision must be made as to how often the Target will be revised and in what way it will be revised. If the Target was intended to reflect what was needed to achieve savings in one year, the same Target may be too high or too low to accurately reflect continued savings in a subsequent year. For example, if per-patient spending would have been expected to increase in the next year if the APM had not existed (e.g., simply through higher prices to cover inflation in the costs of materials and wages), then some increase in spending in the APM could still reflect savings, and a corresponding increase in the Target would be appropriate. On the other hand, if per-patient spending would have decreased in the next year in the absence of the APM (e.g., because a new treatment becomes available that is equally effective but costs less), then the Target would need to be lower to reflect that.

Over time, however, it may be necessary not just to update the Target amounts, but to change the methodology for setting Benchmarks or to move to a different approach to setting Targets. For example, if Prior Performance Benchmarks are used, the initial Target Change will presumably be designed to achieve some reduction in utilization or spending, but once that reduction has been achieved, the Target Change would need to change to zero, or some other method of setting Benchmarks or Targets would be needed.

Example: In the Comprehensive Care for Joint Replacement APM, the Targets are based on a blend of the previous level of spending for the individual provider and the average level of spending in the region where the provider is located. In the first two years of the program, the Targets were weighted primarily toward the individual provider's prior performance, but in the third and fourth years, the Targets are weighted more heavily toward the regional average and in the fifth year, the Targets are based solely on the regional average.¹⁴³ (In effect, the program transitions from using a Prior Performance Benchmark to a Competitive Benchmark.)

If Comparison Group Benchmarks are used initially, it will become increasingly difficult to continue using them if a large proportion of eligible providers and patients are participating in the APM, since there will be a smaller number of non-participants available to include in a comparison group and the non-participants may not really be comparable to those who are participating. Similarly, it may become more difficult to define a meaningful Counterfactual Benchmark if there are no longer adequate current data about utilization or spending for providers or patients that are not participating in the APM to enable judgments to be made about “what utilization or spending would have been in the absence of the APM.” Consequently, if an APM is successful and is used by most or all providers, it will likely need to transition to using Evidence-Based Targets, Competitive Targets, or some other approach.

TABLE 9
ALTERNATIVE WAYS OF SETTING TARGETS FOR UTILIZATION/SPENDING

	Strengths	Weaknesses	Examples
Patient-Level Target	Ensures that spending is appropriate for each patient	Not always possible to specify services individual patients should receive	No medications or tests ordered that are inconsistent with evidence-based guidelines
Population-Level Target	Does not require that spending be reduced for every patient	Allows spending to increase for some patients as long as spending is reduced for others	
A. Evidence-Based Target	Target is known to be achievable without adverse effects	Evidence may not be available for innovative approaches or the types of patients being treated	Average per-patient spending on medications used to achieve good outcomes for the condition in controlled trials
B. Competitive Target	Encourages innovation in lower-cost methods of care	Requires availability of multiple providers who can and will compete based on spending	Lowest average per-patient spending on condition-related medications by other APM providers who achieved good outcomes for similar patients
C. Benchmark-Based Target	Helps ensure Target spending is at or below levels expected under current payment system	Requires a way of determining what spending would have been in the absence of the APM	
1. Benchmark Definition			
a. Prior Performance Benchmark			
<i>Prior Performance for Same Patients</i>	Allows the Target for the provider to be based on the unique needs of the patients being treated	Cannot be used for new acute conditions; Patients with chronic conditions may have been undertreated previously or their needs may have increased	Per-patient spending on medications to treat the targeted health condition in the same patients during the prior year
<i>Prior Performance for Similar Patients of Same Provider</i>	Allows a provider-specific Target to be used for providers treating acute conditions	Will result in higher Targets for providers who have had higher spending in the past; Current patients may have different characteristics; Evidence about treatment may have changed	Average per-patient spending on medications ordered by the APM provider to treat the targeted health condition in similar patients during the prior year
b. Comparison Group Benchmark	Enables the Target to be equal to or lower than what has been achieved in the past for similar patients	Requires a comparison group that is similar in key respects	Average per-patient spending on condition-related medications for similar patients in the previous year treated by providers not participating in the APM
c. Counterfactual Benchmark	Allows the Target to be based on unique characteristics of patients and expected changes over time	Many assumptions must be made about what services would have been delivered in the absence of the APM	Estimated per-patient spending on condition-related medications in the current year based on patient characteristics

CONTINUED ON NEXT PAGE

TABLE 9 (CONTINUED)
ALTERNATIVE WAYS OF SETTING TARGETS FOR UTILIZATION/SPENDING

	Strengths	Weaknesses	Examples
Population-Level Target			
C. Benchmark-Based Target			
2. Target Change			
a. Min/Max Change Needed for Success	Directly ensures net savings by tying savings goal to amount of added spending	May achieve less savings than is possible	Reduction in medication spending needed to offset payments for new services delivered to patients
b. Comparison Group Change	Ensures the Target Change will be equal to or greater than what has been achieved for similar patients	Requires identification of providers not participating in the APM who have similar patients and similar environments	Reduction in per-patient spending on condition-related medications achieved for similar patients by providers not participating in the APM
c. Statistically Significant Change	Ensures that the Target Change cannot be achieved through random variation alone	Biased against providers with small numbers of patients and services with high variability across patients	Reduction in per-patient spending on condition-related medications that is different from zero by a statistically significant amount
d. Desired Level of Change	Ensures that the APM achieves sufficient savings to justify the effort in implementing it	May not be realistic to achieve, particularly for providers who already have low levels of spending	Amount of savings on medications desired by payers or patients

3. Assessing Performance on Utilization and Spending

Unfortunately, assessing the performance of an APM participant on a spending or utilization measure is usually not just a straightforward matter of comparing the measure with the Target. The assessment methodology needs to determine the extent to which any difference between the measure and the Target was due to the APM participant's performance vs. other factors.

a. Causes of Uncertainty in Assessing Performance

There are five types of reasons why actual utilization/spending on a particular measure may exceed or fall short of the APM Target:

1. **Performance Success or Failure.** Whether the Target is reached will depend on whether the APM participant takes the necessary actions to reduce, or to prevent an increase in, utilization/spending. For example, if the APM was designed to reduce avoidable hospitalizations for patients with a chronic disease, there could be inadequate savings on hospitalizations if the care managers supported by the APM failed to adequately address the relevant aspects of the patients' health or behavior in a way that would reduce their need for hospitalization.
2. **Errors in Calculation or Measurement.** The Benchmark or the measure of actual utilization/spending could be computed inaccurately. For example, if the patients involved in the APM had comorbidities that were not accurately recorded, or if a hospitalization was inaccurately classified as "avoidable," then a comparison of the measured utilization/spending to the Target would not accurately reflect the APM participant's true performance.
3. **Effects of Uncontrollable Factors.** In order to properly assess performance, both the utilization/spending measure and the Target for the measure have to be correctly and completely adjusted for known factors that affect utilization/spending but are not controllable by the APM participant. For example, if the goal of the APM was to reduce avoidable hospitalizations, but the Benchmark calculation failed to adjust for important patient characteristics that are known to have a large effect on the rate of avoidable hospitalizations, then an APM participant's actual utilization/spending could be higher or lower due to selective participation or non-participation of patients with those characteristics, not because of any impacts of the services supported by the APM.¹⁴⁴
4. **Effects of Rare or Unpredictable Events.** Some circumstances that affect utilization and spending will occur infrequently and unpredictably, e.g., a major flu outbreak in a community, a patient with a rare variant of a disease, or a patient experiencing an unusual confluence of problematic circumstances each of which requires additional services to address. It may be possible to predict the effect these circumstances will have on spending but not whether the circumstances themselves will occur during the period of time in which performance is being measured.

5. **Effects of Random (Unexplained) Variation.** Finally, there are differences among patients in their needs for services and changes in their needs over time that occur for reasons that cannot be explained or predicted, much less controlled by the APM participant.

The extent to which the difference between the Target and the APM participant's utilization/spending truly reflects the participant's success or failure in changing utilization or spending depends on the relative size of the first factor versus the other four. It would be inappropriate to credit the APM participant with success in reducing spending if most or all of the difference in spending was due solely to an error in measurement or a failure to properly adjust for differences in the patients being treated, and it would be inappropriate to penalize the provider participating in the APM because it failed to reduce spending due solely to rare events or random variation in spending. These same factors can affect both the Target amount and the APM participant's actual utilization/spending. For example, if a Benchmark is used in setting the Target, the Benchmark amount may be too high or too low because of a failure to adjust for differences between the patients used for calculating the Benchmark and the patients whom the APM Participant is treating.

Example: The rate of new lung cancer diagnoses is about 300 per 100,000 Medicare beneficiaries, which means that an ACO in the Medicare Shared Savings Program that has 10,000 assigned beneficiaries might expect to see an average of 30 new lung cancer cases each year. However, the number could easily vary significantly from year to year. Since treatment for lung cancer can cost as much as \$100,000 per patient, an ACO that had 35 lung cancer cases in a year could spend as much as \$1,000,000 more than an ACO with only 25 such cases, solely due to the difference in the number of cancer cases. That alone would make the first ACO's overall average spending per beneficiary 1% higher than the other's.

The effects of the second factor can potentially be reduced through more careful collection of data and calculation of measures. In principle, the effects of the third factor can be reduced through improved risk adjustment methodologies, but despite years of research, most risk adjustment methodologies only explain a small fraction of the variation in services that patients will need. As for the fourth factor, most risk adjustment systems do not adjust for rarely-occurring factors because they are rare. Although special adjustments could be made after such events occur, the rarer the event, the more difficult it will be to know what adjustment is appropriate. In addition, if special procedures or additional costs are required to identify or verify the presence of a rare event, the less likely it will be that such events are recognized when they do occur.¹⁴⁵

In most cases, the biggest challenge in accurately assessing performance will be random variation. There is a large amount of "random" (i.e., unexplained) variation in most measures of utilization and spending, and this will

create considerable uncertainty as to whether a difference between the measured level of utilization/spending and the Target represents an actual change in utilization/spending and whether the change is attributable to actions by the providers or to other, unknown factors.

b. Balancing Type I vs. Type II Errors in Addressing Random Variation

Many APMs have recognized the importance of random variation in assessing performance, but the focus has primarily been on avoiding an erroneous declaration that an APM participant has been successful when the target level was achieved solely because of random variation. In statistics, these are known as “Type I” errors. It is impossible to guarantee that there will be no Type I errors in an APM, and so the frequency of Type I errors that is considered tolerable is used to set the significance level in statistical tests. For example, if the performance assessment methodology determines that the target level was achieved with “95% confidence” or at the “5% significance level,” it means that if the same situation occurred on multiple occasions, spending would have achieved the target level solely due to random variation only 5% of the time.

However, there is a second kind of error in statistical analysis called “Type II” error. A Type II error occurs when a determination is made that the target level was *not* achieved even though the APM participant actually *did* achieve the Target. The likelihood of a Type II error is directly and inversely related to the level of confidence chosen regarding Type I errors. For example, the more certain one wants to be that the Target was not achieved solely due to random variation, the more likely one is to *wrongly* believe a spending change was due to random variation when it actually represented the effects of the APM participant’s actions, and vice versa.

In addition, the likelihood of a Type II error also depends on how big the actual change in utilization/spending is compared to the Target. If the APM participant’s performance is much better than the Target, then the chance of wrongly attributing the performance to random variation will be much lower.

The practical effects of Type I and Type II errors are that:

- It will generally not be sufficient for the APM participant to achieve a level of utilization or spending *equal* to the Target; the APM participant will need to reduce utilization/spending *below* the Target level by some amount in order to provide confidence that the Target had not been reached solely due to random variation. In effect, the “real” Target will need to be lower than a Target determined through the options described earlier.
- The higher the desired level of confidence that the Target has truly been achieved, the more likely it is

that participants will be incorrectly determined to have failed when they have actually succeeded.

Because of this tradeoff between the two error rates, there is no “right” way to set the significance threshold for assessing performance on a measure of utilization or spending. Payers will understandably want the highest level of confidence possible that an APM participant has truly achieved the Target level of utilization or spending, but the APM participant will want the greatest assurance possible that they will not be penalized for poor performance when they have actually achieved the results desired. If payers demand too high a confidence level, providers will be unwilling to participate in the APM, but if payers do not have confidence that the necessary performance levels will be achieved, they may be unwilling to implement the APM or continue using it.

Example: In the Medicare Shared Savings Program, in order to qualify for a shared savings payment, an ACO with fewer assigned patients is required to achieve a higher Minimum Savings Rate than a larger ACO because the smaller the number of patients assigned to the ACO, the lower the statistical confidence will be that any savings were not due to random effects. However, the Minimum Savings Rates for small ACOs are lower than they would otherwise be because CMS used a lower level of statistical confidence in setting the Minimum Savings Rates for smaller ACOs.¹⁴⁶

The ideal situation occurs when APM participants can easily perform better than the Target needed for success of the APM. In that case, the rate of both types of errors will be lower, i.e., the high performance level that effective APM participants will achieve will provide a high level of confidence that the Target has been met or exceeded. Conversely, if an APM participant achieves

the Target but does not exceed it, that participant should appropriately be viewed as less successful due to the greater probability that the achievement was due to luck.

Factors Affecting the Magnitude of Errors in Performance Assessment

There are several factors that increase the probability of errors in performance assessment due to random variation:

- **Smaller Numbers of Patients.** If only a small number of patients is eligible for the APM, or if a particular APM participant only treats a small number of patients, then the effects of random variation on utilization/spending will be higher. An additional problem for APM participants with small numbers of patients is that more situations will need to be treated as “rare events” than with larger providers, because a low frequency event may not occur at all in a given year if there are only a small number of patients.

Example: Assume that patients with a particular condition can be treated for \$1,000 on average, but a less common variation of the condition occurs in 1 out of 1,000 patients, and when that occurs, the treatment cost for the patient jumps to \$10,000. If an APM participant treats 10,000 patients per year, it could expect that about ten of the patients each year would have the infrequent variation, and the overall average cost of treating all of its patients would be \$1,009 ($9990 \text{ patients} \times \$1,000 + 10 \text{ patients} \times \$10,000$)/10,000). Even if that APM participant had one more or one fewer patient with the infrequent but expensive variation during the year, it would not affect the overall average spending by very much. (For example, if 11 patients happen to have the condition in a particular year, the average cost that year would be \$1,010 ($9989 \times \$1,000 + 11 \times \$10,000$)/10,000).

However, if a small provider participating in the APM only treats an average of 100 patients for the condition during the course of the year, then in most years, none of the patients will have the rare condition and the average cost of treating the 100 patients in most years will be \$1,000 (nearly 1% less than the average cost for the APM participant with 10,000 patients). In contrast, in a year in which one of the 100 patients had the rare condition, the average cost of treating all of the patients would be \$1,090 ($99 \text{ patients} \times \$1,000 + 1 \text{ patient} \times \$10,000$)/100), which is 8% higher than the average for the larger participant.¹⁴⁷

- **Short performance periods.** Over time, random variations would be expected to average out, but the shorter the period of time in which performance is being measured, the less opportunity there will be for that averaging to occur. Moreover, during a shorter period of time, fewer patients will be included, which will also increase the magnitude of random variation relative to true changes in utilization and spending.

In the previous example, if spending were calculated every month rather than over a year (and if the treatments for patients required less than a month to complete), the average cost of treatment in most months would be \$1,000, but in the month in which the patient with the rare condition was treated, the average cost would be over \$2,000.¹⁴⁸

- **Diverse groups of patients.** The greater the diversity of patients being treated as part of the APM, the more likely it is there will be differences among them that affect spending and the fewer patients there will be in any one of the categories used to stratify for variation. For example, a primary care practice will see a much more diverse array of patient conditions than would a similarly-sized single-specialty practice. In addition, standard methodologies for evaluating statistical significance are based on a series of assumptions about the variance in the underlying patient populations, and those assumptions may not be appropriate for the types of patients and services being analyzed. For example, typical methodologies assume that variables follow a normal distribution, but the actual distribution of patient spending is generally highly skewed, and

this causes confidence intervals to be inappropriately narrow when calculated using standard statistical techniques.¹⁴⁹

- **Composite measures of utilization/spending.** As explained in Section VI.B.1.b, if performance is assessed using a total cost of care measure or other composite measure that includes both the types of services the APM participant would be expected to change and types of services the APM could not affect, a statistically significant change in spending on a specific service may only produce a statistically insignificant change in spending on the large composite.¹⁵⁰

There are ways that an APM can be designed to reduce the effects of these factors, but these design features can be problematic in other ways. If only large provider organizations are permitted to participate in an APM, the impacts of random variation on the APM's accountability component will be reduced, but the overall impact of the APM will also be reduced because fewer providers and thereby fewer patients in total will be able to participate. If spending is measured over a longer period of time, the effects of random variation will be reduced, but there will be greater cash flow complications for providers, payers, and patients because of the longer delays between when services are delivered and when success on spending can be determined. If utilization/spending is measured only for the specific services targeted by the APM, the errors in assessing performance on those services will be lower, but there will be greater uncertainty about whether the APM is causing increases in other services in unexpected ways.

c. Bayesian Methods

An important way of reducing the likelihood of both Type I and Type II errors, particularly for small providers, is to incorporate additional information about likely levels of performance under the APM. For example:

- if an APM participant has a small number of patients, random variation can make it hard to determine what the APM participant's true performance level is. If a participant's performance in the current year was positive but failed to achieve the level of statistical significance required, rather than calling that a failure, payers and patients could look at the participant's performance in prior years. If the APM participant successfully reduced utilization while maintaining quality in previous years, that should make payers and patients more confident that current performance was also likely to have been successful and that it was random variation that made performance appear worse.
- if a newly-created APM is designed to support a change in care delivery that has been shown in pilot projects to have reduced spending and improved quality, there should be greater confidence that initial reductions in spending under the APM represent "true" savings even if too few patients have been treated to reach a particular statistical significance level.

Traditional "frequentist" statistical methods have no way of incorporating this kind of prior knowledge, and they simply calculate statistical significance levels based on

Minimum Savings Rates in the Medicare Shared Savings Program

In Section 1899 of the Social Security Act, which established the Medicare Shared Savings Program, Congress specified that an Accountable Care Organization (ACO) could receive a shared savings payment “only if the estimated average per capita Medicare expenditures ... is at least the percent specified by the Secretary below the applicable benchmark...The Secretary shall determine the appropriate percent ... to account for normal variation in expenditures ... based upon the number of Medicare fee-for-service beneficiaries assigned to an ACO ...”

When it issued regulations to implement the MSSP, CMS defined the “Minimum Savings Rate” (MSR) that an ACO would need to achieve in order to qualify for a shared savings bonus payment. The MSR is the minimum percentage difference between actual total per-beneficiary spending and the ACO’s benchmark spending level that the ACO would have to achieve in order for CMS to conclude that the ACO had achieved savings. (This is similar to the “Statistically Significant Change” option for defining a Target Change described in Section VI.B.2.) For ACOs with “two-sided risk,” CMS also defined the “Minimum Loss Rate” (MLR) that would trigger a penalty (i.e., a payment to CMS by the ACO).

Congress seems to have recognized that errors could be more likely for an ACO with a smaller number of assigned patients, which is why the law says the percentages should be based on the number of assigned beneficiaries. In addition, CMS recognized there was a tradeoff between Type I and Type II errors when it chose the MSR and MLR levels for different sizes of ACOs, since the minimum savings rates were based on different levels of confidence about the effects of random variation (i.e., different Type I error rates):

- In Track 1 of MSSP (where the ACO can receive a shared savings bonus if spending decreases but the ACO does not have to pay a penalty if spending increases), the MSR was set at 2.2% for ACOs that have 50,000 patients. This was based on a confidence level of 99% that a shared savings bonus would only be awarded if savings were actually achieved.*
- For Track 1 ACOs with 20,000 patients, the MSR was set at 2.5%, based on a confidence level of 95%.*
- For Track 1 ACOs with only 5,000 patients, the MSR was set at 3.9%, based on a confidence level of 90%. (The MSR would have needed to be even higher if CMS had used the 95% or 99% confidence level it used for larger ACOs.)*
- For Track 2 ACOs (which pay penalties if spending increases as well as receive bonuses if spending decreases), the MSR and MLR were set at 2% regardless of the number of patients assigned.*

MSRs and MLRs were not established for smaller numbers of patients because Congress had limited participation in the Medicare Shared Savings Program to ACOs with at least 5,000 patients. This was presumably based on analyses indicating that the variability in total spending per patient with smaller numbers of patients would have caused high rates of errors in determining whether savings had occurred.

Analyses by Derek DeLia and colleagues¹⁵¹ have shown that the low rates of Type I errors used by CMS in setting the MSR and MLR result in relatively high probabilities of Type II errors, particularly for smaller ACOs. For example, in a Track 1 ACO with 5,000 assigned beneficiaries, if the ACO reduced Medicare spending by 4%, there would be a 49% probability that the ACO would not receive a shared savings bonus (because random variation could push the measured spending below the 3.9% MSR about half of the time). If the ACO with 5,000 assigned beneficiaries was in the two-sided risk model and reduced spending by 4%, there would only be a 26% probability that it would not receive a bonus (because the MSR is only 2% in the two-sided model), but there would be a 2% probability that random variation could cause it to appear that the ACO had actually increased spending by enough to trigger a penalty. Even in an ACO with 50,000 assigned beneficiaries that achieved savings of 3% (i.e., 50% more than the minimum savings rate), there would be a 20% probability that the ACO would receive no shared savings bonus. DeLia has also shown that the true Type I and Type II error rates are likely higher than what is estimated by the methodology used by CMS, because there is also random variation in the historical spending data used to set the ACO’s benchmark as well as potential errors in the trend factor used in calculating the benchmark.¹⁵²

the data from the current performance period. However, Bayesian statistical analysis methods (e.g., hierarchical generalized linear models) provide formal ways of incorporating additional information in ways that can create greater certainty in the evaluation of performance. In Bayesian analysis, one considers not just data on a provider's actual performance on spending measures but also the "prior probability" that various results would occur on the measures. If there are enough actual data to confidently determine what has happened, then those data determine the result, but if the actual data are limited (e.g., because individual providers only have a small number of patients in the APM in any given year), then the prior probabilities are combined with the actual data in order to create the equivalent of a larger sample size. If there is little or no information available to determine the prior probabilities, then they contribute relatively little to the conclusions, but if enough information is available to create "strong priors," then that information complements what is available in the current data. CMS uses Bayesian methods in determining how to evaluate hospital mortality, and this approach has been endorsed by nationally-recognized experts in statistics and performance measurements.¹⁵³

In addition, for providers with small numbers of patients, Bayesian analysis provides a rigorous way to update the evaluation of performance over time. When the APM is first implemented, there may not be sufficient data on a small provider's own patients to confidently evaluate the provider's performance. Over time, Bayesian methods can use the evaluations of the provider's performance in prior years to create the equivalent of a larger sample size for evaluating their performance in the current year. If performance has been consistently high in the past, it increases the confidence that high performance currently is not due to random variation, and vice versa.

The additional complexity of Bayesian methods often deters their use. Many people have difficulty understanding traditional frequentist statistics, and Bayesian statistics does not lend itself to simple formulas for calculating significance levels. However, the solutions that are typically pursued may also be undesirable.

Example: Participation in the Medicare Shared Savings Program is statutorily limited to groups of providers who have more than 5,000 Medicare patients because the change in spending required to statistically "prove" that the change was not due to random effects would be unrealistically high for smaller ACOs. However, the majority of counties in the United States don't have 5,000 Medicare beneficiaries living in them; this means that providers in these small and rural counties are precluded from participating in the program unless they form a partnership with other communities to create a larger ACO. In some cases, ACOs have been formed by combining small counties in completely different parts of the country solely to achieve the minimum number of patients needed to qualify as an ACO or to qualify for a lower Minimum Savings Rate. The result is that an ACO with "good performance" on savings could consist of one community where spending increased (or quality decreased) and a second community where savings were achieved. Such an ACO could be determined by CMS to be successful, even though the pa-

tients in the first community would not have received the benefits that the ACO was intended to produce. No real increase in certainty about whether savings occurred is achieved by simply combining the results from separate communities.

d. Size and Significance Are Both Important

Although it is important to recognize the impacts of random variation and to try to avoid drawing incorrect conclusions because of it, an excessive focus on statistical significance can lose sight of the real goal of APMs – reducing spending and improving quality.

Example: Although a 95% confidence level about the amount of savings is preferable to 80% confidence, if APM #1 is determined to have decreased spending by \$1,000 per patient with 80% confidence and APM #2 decreased spending by only \$100 but with 95% confidence, a rational decision-maker should choose APM #1. (The expected savings in APM #1 is \$800 per patient, whereas the expected savings in APM #2 is only \$95.) Unfortunately, the actuarial approach typically used to evaluate APMs today might well declare that APM #1 had failed and that APM #2 had succeeded.

In most cases, the choice is not which of two APMs to implement, but a choice between implementing an APM or continuing with the fee for service system. The higher the level of confidence about savings a payer demands from the APM, the more likely it is that the APM will be rejected in favor of continuing with the standard payment system. A more realistic comparison would treat fee-for-service as a different type of APM – one where there is no requirement for savings at all and a high probability (based on past experience) of an increase in spending. In that comparison, an APM that has "only" an 80% probability of achieving savings would look highly desirable rather than being declared a failure.

Moreover, the significance levels commonly used in evaluations are merely matters of convention and are inherently arbitrary. The standard 95% confidence level is only slightly better than 94% confidence, yet using rigid evaluation criteria would cause an APM that achieved a large amount of savings with 94% confidence to be declared a failure but an APM that achieved a small amount of savings with 95% confidence would be called a success. There are a variety of assumptions that must be made in statistical analysis in order to calculate confidence levels, and those assumptions also introduce error into the calculations, so it is even more problematic to conclude whether an APM is successful or not based solely on whether a change in spending reached a particular level of significance.

From the perspective of patients, both the size of the impact and the certainty of that impact are important. While achieving a larger amount of savings on average will be more desirable for most patients, a lower probability of achieving savings means that an individual patient cannot be sure what benefit they will achieve.

Consequently, a good performance assessment methodology should consider both the size and the certainty of a provider's performance in determining success or failure despite the inherent difficulty of doing so.

4. Making Performance-Based Adjustments to Payments

Once the spending measures, targets, and methods of assessing performance are defined, the final step is defining the mechanism of accountability, i.e., the actions that will be taken if the actual performance on one or more of the measures is determined to have fallen short of the target level.

There are five basic options for accountability:

1. Penalties or bonuses based on whether Targets are achieved, in addition to payments for the services delivered to patients;
2. Paying for services to patients only when Targets are achieved, i.e., “outcome-based payments;”
3. Bundled/warranted payments that require APM participants to pay for services or spending that are included in performance measures;
4. Terminating a provider’s participation in the APM if the provider fails to achieve Targets; and
5. Terminating the APM altogether if providers fail to achieve Targets.

Option 1: Penalties/Bonuses in Addition to Service-Based Payments

Under this option, the healthcare provider that is participating in the APM is paid for delivering the desirable services using whatever methodology is defined in Component #1 but the APM participant is required to pay a penalty if the Targets on one or more utilization/spending measures are not achieved.

This option requires a methodology for determining the magnitude of the penalty that will be imposed if the Target for a utilization/spending measure is not achieved. The larger the penalty, the greater the financial incentive for the APM participant to achieve the desired result, but the greater the potential financial problems the participant could face, particularly if the penalty represents a large proportion of the participant’s revenue.¹⁵⁴

a. Patient-Level Targets vs. Population-Level Targets

Penalties can be used with either Patient-Level Targets or Population-Level Targets.

- If a Patient-Level Target is not met for an individual patient, the provider would pay a penalty for that patient, e.g., by charging the patient less or refunding all or part of the payments that the patient or their payer had already made.
- If a Population-Level Target is not met, the provider would either pay a penalty to the payer for a group of patients or would pay each individual patient a share of the overall penalty.

An important weakness of using Population-Level Targets under Option 1 is if the provider participating in the APM achieves the Target level of utilization or spending, the provider would pay no penalty, even though spending for a subset of patients could be higher than it would have been otherwise or the patient may have received

undesirable services that the APM was supposed to avoid. If those patients are paying for their own care, or their health insurance requires cost-sharing for services (whether it be in the form of a deductible, co-insurance, or a copayment), those patients would pay more under the APM rather than less. From the patient’s perspective, this is the opposite of how a value-based payment should work – the patient with a bad outcome shouldn’t be paying more than the patient with a good outcome. Patient-Level Targets can avoid this problem.

b. Making the Penalty Proportional to Performance on the Measure

The impact on payers and patients of a provider’s failure to achieve a Target will often vary based on the amount by which utilization/spending differs from the Target, so rather than imposing a fixed penalty for failure to meet a Target, the penalties could be scaled based on the amount of the shortfall. For example, if a home nursing service was expected to reduce the rate of avoidable hospitalizations by 10%, but the rate was only reduced by 5%, then the penalty imposed could be lower than if there was no reduction in avoidable hospitalizations at all. If the rate of avoidable hospitalizations actually increased, the penalty could be even higher than it would have been for no reduction at all. The difference in the penalty could be directly proportional to the actual performance (e.g., if the rate of hospitalizations was reduced by 5% rather than 10%, the penalty could be 50% of what it would be if there were no reduction at all), or a more complex formula could be defined.

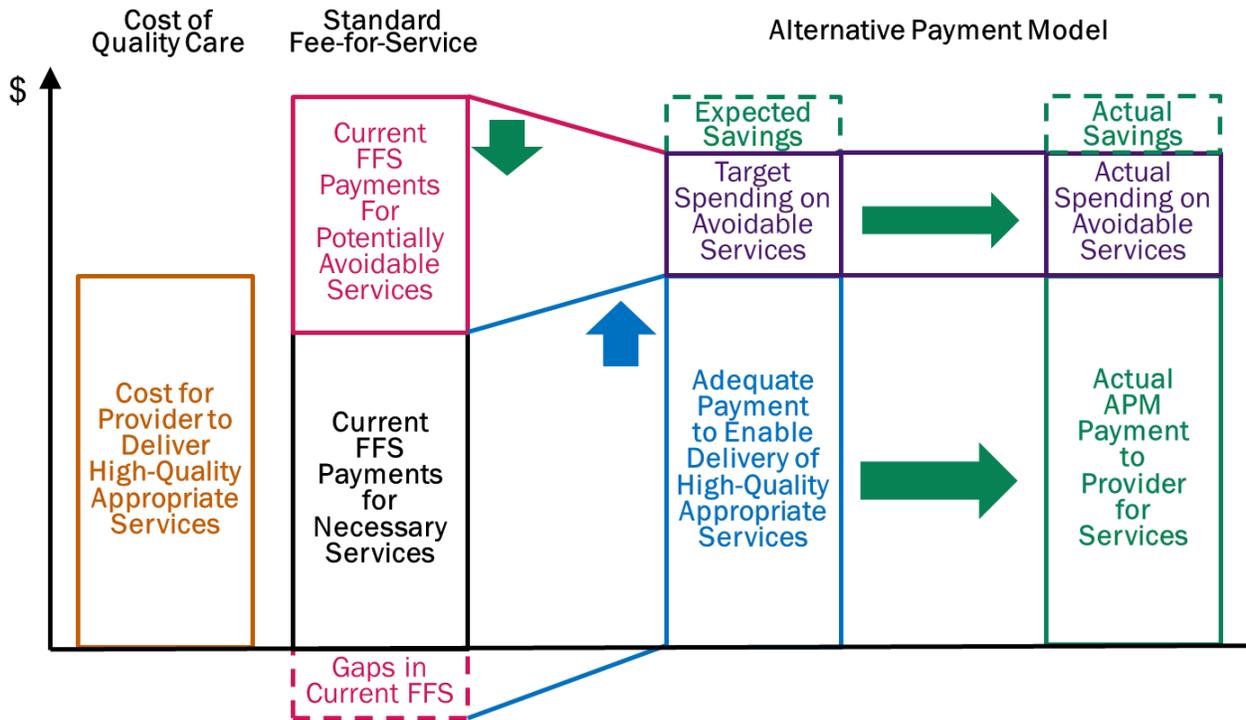
This approach is helpful when it is difficult to predict the exact impact of a service or where there is random variation in the desired result, since it avoids imposing a large penalty if the result falls just slightly below the target change.

c. Determining the Amount of the Penalty

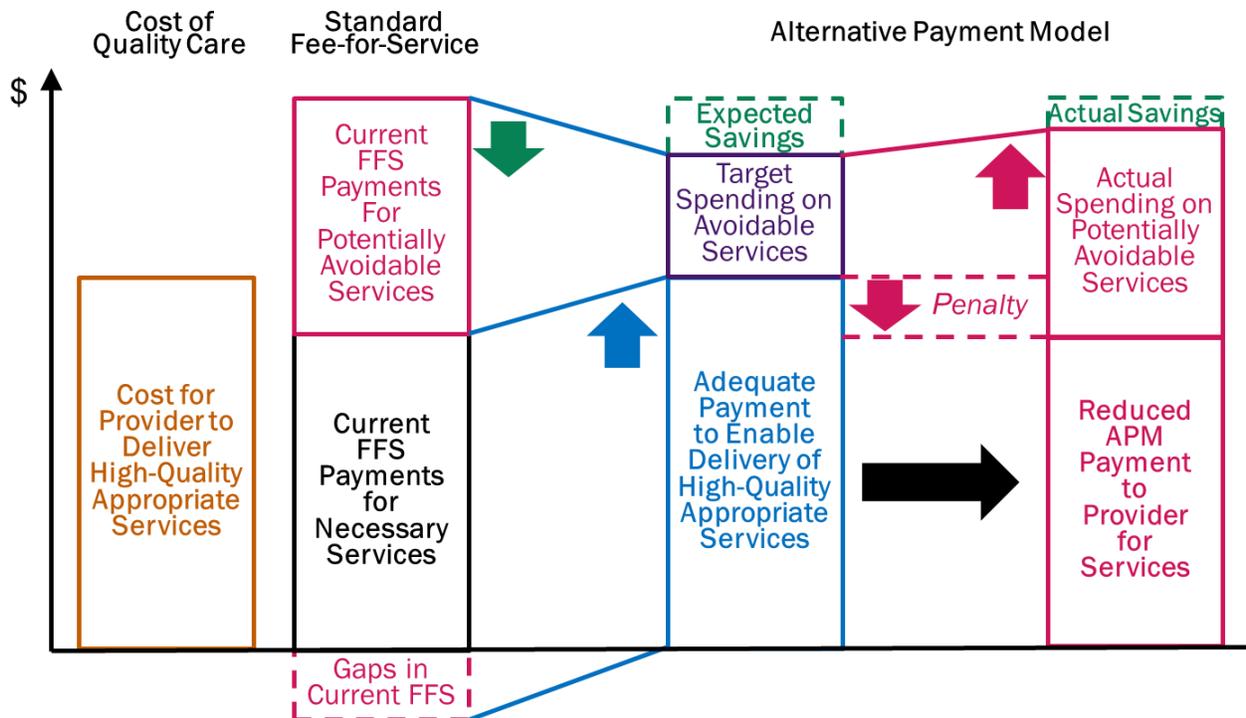
It is not enough to decide the relative magnitudes of penalties; the absolute amount of the penalty must also be determined. There are many different ways that the actual amount of the penalty can be determined. Two options are:

1. **Penalty Amount Based on the Amount of Payment for Planned Services.** The amount of the penalty can be tied to the amount of payment the APM participant is receiving for the services that were intended to affect the utilization/spending measure. In the example above of an APM for a home nursing service, if there is no reduction in the rate of avoidable hospitalizations, the APM participant could be required to repay a portion of the payment for the home nursing service, and if the rate of avoidable hospitalizations is only reduced by half of the target amount, the APM participant could be required to repay an amount half as much as it would have to pay if there was no reduction in hospitalizations at all.
2. **Penalty Amount Based on the Target Change in Spending.** The amount of the penalty can be tied to the amount of savings that was expected to be

**FIGURE 5
FULL PAYMENT UNDER COMPONENT #1 WHEN TARGET IS ACHIEVED**



**FIGURE 6
PENALTY (LOWER COMPONENT #1 PAYMENT) FOR FAILURE TO ACHIEVE TARGET**



EXAMPLE OF AN APM WITH A PERFORMANCE PENALTY

Assume that a nurse is hired to make home visits to patients with a moderately severe chronic disease in order to reduce avoidable admissions to the hospital. The annual cost associated with the nurse's compensation, travel costs, etc. is \$100,000. If the nurse can make regular home visits to 100 patients, the average cost per patient per year would be \$1,000. Assume that 25% of the provider's patients with the disease have been admitted to the hospital during the course of the year, that a hospital admission costs \$10,000, and that having the nurse make home visits is expected to reduce the hospitalization rate to 10%. As shown in the table, if the home visits achieve the expected results, the total savings for the payer would be \$150,000 (100 patients x 15% avoided admissions x \$10,000), which would be more than the \$100,000 needed to pay for the cost of the nurse visits. The payer could pay for the visits on a per-visit basis, a per-patient basis, or some combination, as described in Section VI.A, with the payment amount defined based on the expected cost of delivering the service. The Target performance could be set equal to 10%, based on a Prior Performance Benchmark of 25% and a Target Change of 15%.

In order to ensure the home visits are delivered in a way that reduces the rates of hospitalizations, the APM could create a penalty for poor performance by reducing the provider's payment by 1% for each 1% that the actual hospitalization rate is above the 10% Target. As shown in the table, if the actual hospitalization rate was 15%, the provider would experience a loss of \$5,000 on the nursing services, but the payer would still spend less in total than previously.

	FFS	Penalty-Based APM	
	Current Performance	Target Performance	Failure to Meet Target
Patients			
Total # of Patients	100	100	100
% Patients Hospitalized	25%	10%	15%
# Patients Avoiding Hospitalization	75	90	85
Payment Per Patient for Nurse Home Visits		\$1,000	\$1,000
# Patients Hospitalized	25	10	15
Payment Per Patient for Nurse Home Visits		\$1,000	\$1,000
Provider Revenue/Cost			
Revenue from Payments for Nurse Home Visits		\$100,000	\$100,000
Penalty = % Difference Between Performance & Target		\$0	(\$5,000)
Cost of Nurse Making Home Visits		(\$100,000)	(\$100,000)
Profit/Loss for Provider for Nurse Home Visits		\$0	(\$5,000)
Payer Spending			
Spending on Hospitalizations (\$10,000 per admission)	\$250,000	\$100,000	\$150,000
Payments for Nurse Home Visits		\$100,000	\$100,000
Bonus/Penalty for Performance Compared to Target		\$0	(\$5,000)
Total Payer Spending	\$250,000	\$200,000	\$245,000
% Change in Spending		-20%	-2%

achieved in the measure of spending. In the example, the APM participant could be required to pay a penalty proportional to the amount the payer has to spend on the hospitalizations that were not avoided.

Approach #1 assures the payer (including the patient) that it will not be making the full amount of the new or modified payments for services under the APM unless the services are effective in achieving the goal. This is, in effect, a full or partial money-back guarantee on the payments the APM participant receives.

In contrast, Approach #2 assures the payer that it will receive all or part of the expected savings regardless of whether the Target Change in utilization or spending actually occurs. This is the equivalent of a performance warranty, since the APM participant would still be paid for delivering the desirable services, but it would also be responsible for paying for all or part of the undesirable services if they occur.

Relative Financial Risk Under the Different Approaches

Both Approach #1 and Approach #2 create “financial risk” for the APM participant. Assuming that the amounts of the payments for the desirable services are based on the costs the APM participant would be expected to incur in delivering those services, then a reduction in the payment for those services under Approach #1 could mean that the APM participant’s revenues for the services would be less than needed to cover its costs. Under Approach #2, there is no change in the APM participant’s payments for the services, but the APM participant’s costs increase, because of the need to pay the penalties in addition to the costs incurred in delivering the services.

The key difference between Approach #1 and Approach #2 is that if the Target was intended to achieve net savings for the payer – i.e., the spending reduction from achieving the Target would be greater than the amount of new payments for the desirable services – then the financial risk for the provider under Approach #2 could be much larger than under Approach #1. Depending on the amount of spending associated with the performance measure, the provider participating in the APM could conceivably be responsible for paying the payer more than the total revenues the APM participant receives for all services, not just the service delivered as part of the APM. For example, the amount that a payer spends on a single hospitalization will typically be 100 times or more the amount it pays for an office visit with a physician, so if a physician is responsible for paying a penalty based on even a small number of hospitalizations, paying that penalty could require all of the revenue the physician receives for all of his or her services. Depending on the magnitude of the financial risk, a provider participating in an APM using Approach #2 could be subject to insurance regulation.

Not surprisingly, the fact that Approach #2 involves greater financial risk for the provider means that Approach #2 has less financial risk for the payer than Approach #1. However, whereas either Approach #1 or Approach #2 increases the financial risk for the provider compared to what it would have been in the absence of

the APM, only Approach #2 will reduce the payer’s current financial risk. This is because under the current payment system, the payer is at full financial risk for any spending associated with the undesirable services that are being measured. Under Approach #2, the provider either reduces that spending or takes responsibility for paying a portion of the spending that is not reduced, so the payer’s risk is reduced. Under Approach #1, if the provider does not reduce the avoidable spending, the payer may still have to pay more for a portion of the unsuccessful services delivered by the provider as well as paying for the services that were not avoided.

Other Approaches

The penalty could be determined in some way that is not tied directly to the amount of either the desirable or undesirable spending, or the amount of the penalty calculated under Approach #1 or Approach #2 could be varied in some way based on other factors. For example, penalty amounts could vary based on the level of statistical confidence that the target was not reached, or the penalties could be based on whether the amount by which performance fell short of the target had improved or worsened compared to the prior year. In addition, a combination of Approach #1 and Approach #2 could be used.

Example: The CMMI Comprehensive Primary Care Plus (CPC+) APM uses a variant of Approach #1. A participating primary care practice receives a new monthly payment per patient in addition to fee-for-service payments to support improved primary care, and it is expected to achieve rates of inpatient hospitalization utilization and emergency department utilization that are below a target level. If the practice fails to do so, it has to repay CMS a portion of the additional payments it received.¹⁵⁵

Example: The “downside risk” APMs developed by the Center for Medicare and Medicaid Innovation (CMMI) to date use Approach #2. For example, in the CMS Bundled Payments for Care Improvement – Advanced APM, the provider is required to pay CMS a penalty (labeled the “Repayment Amount”) equal to the amount by which actual spending exceeds the Target Price.¹⁵⁶

Example: A combination of Approach #1 and #2 is what CMS proposed to use in Track 2 of the Oncology Care Model. A participating oncology practice would receive a monthly payment (a “MEOS” payment) to support enhanced services to patients receiving chemotherapy. It would then be accountable for reducing average total Medicare spending on those patients, including the monthly payments, by 2.75%. If the practice failed to achieve this target, it would be responsible for making a payment to Medicare (called the “recoupment”) equal to the amount by which Medicare spending exceeded the target. This means that if total spending other than the monthly payments decreased by only 2.75%, the practice would have to repay the amount of the monthly payments, and if total spending other than the monthly payments decreased by less than 2.75%, the practice would also have to pay for a por-

tion of the amount Medicare spent on other services the patient received.¹⁵⁷

d. Limiting the Financial Impact of the Penalty on the APM Participant

As noted above, the potential for a penalty creates financial risk for an APM participant, and the risk could be very high if the amount of the penalty is proportional to the spending on expensive services such as hospitalizations. Particularly during the initial phases of implementation of an APM, there may be limited understanding of the factors that affect patients' needs for services or the reasons for complications, and so poor performance for some APM participants could be due to differences in the types of patients they are treating rather than their failure to deliver care appropriately.

Payers have promoted the idea that greater financial risk will create a greater "incentive" for APM participants to succeed, but a high level of financial risk could discourage participation in the APM, and large financial penalties could force some APM participants out of business, which would reduce access to care for their patients.

In order to limit the financial impact of penalties on APM participants, three approaches could be used:

1. **Setting maximum amounts on total penalties.** Under this approach, a maximum total penalty amount would be established, and if the APM participant fails to achieve a Population-Level Target, or fails to achieve a Patient-Level Target for a large number of patients, the actual penalty would be the lesser of (a) the amount calculated using one of the approaches described earlier and (b) the maximum amount. Two approaches that have been used for establishing the maximums are:

- a. A percentage of the Target spending level; and/or
- b. A percentage of the APM participant's revenue.

For example, in the CMS Bundled Payments for Care Improvement – Advanced APM, if spending exceeds the target, the maximum amount a provider is expected to repay is 20% of the total of the target spending amounts for all of the episodes.¹⁵⁸

In order for an APM to qualify as an "Advanced APM" under the regulations issued by CMS to implement MACRA, a provider participating in the APM would have to repay Medicare when spending exceeds the Target up to either 3% of the Target amount or 8% of the average revenue of participating providers, whichever is less.¹⁵⁹

2. **"Sharing" risk of higher utilization/spending.** Under this approach, if the APM participant fails to achieve the Target, the penalty amount calculated using one of the approaches described earlier would be reduced by a specific percentage. For example, if the penalty is reduced by 50%, then the APM participant and payer are "sharing" the penalty amount on a 50/50 basis. Two approaches that could be used for establishing the sharing percentage are:

- a. **The APM participant's ability to control the utilization/spending.** If the utilization/spending being measured is completely within the control of the

APM participant, then it may be appropriate to hold the participant 100% responsible for failure to achieve the Target. However, if a measure of total spending on the patients is being used, it would be appropriate to reduce the penalty by more than 50% if the APM participant is only able to affect a small part of total spending.

- b. **The relative size of the APM participant's revenue compared to the Benchmark spending.** The financial risk of a penalty for the APM participant depends on the participant's ability to pay the penalty, which in turn depends on size of the penalty relative to the total amount of revenue the participant receives. In contrast, the financial risk to the payer if the APM participant does not pay the full penalty depends on how big the penalty is compared to the payer's spending on services to the patients in the APM. Consequently, the percentage of the penalty the APM participant would be required to pay could be based on the participant's total revenue relative to the payer's total spending on the patients.

Example: Assume that a primary care practice has 1000 patients with chronic disease, and it currently receives an average of \$500 per patient per year in payments for services it delivers, or a total of \$500,000 in revenue. Assume that the health insurer for the patients spends an average of \$15,000 per year per patient on all services, including hospitalizations that could potentially be avoided, for a total of \$15 million per year in spending on the patients. Assume that the practice participates in an APM that pays it an additional \$1,000 per patient per year to deliver care management services, with a goal of reducing the rate of avoidable hospitalizations for the patients from 25% to 10%, but the rate of avoidable hospitalizations is reduced to only 20%. If the payer spends \$10,000 on each hospital admission, then the 10% of patients whose admissions were not avoided would represent an additional \$1,000,000 in spending for the payer, or 6.7% of the payer's total spending. If the practice was responsible for paying a penalty equal to that full amount, it would represent 67% of the practice's total revenues for the patients. However, if the penalty was set at 10% of the spending on the non-avoided admissions, then the penalty would only be \$100,000 instead of \$1,000,000, which would represent 6.7% of the practice's revenues. Thus, the financial risk of failure for the practice would be equivalent in percentage terms to that of the payer.

3. **Creating a provider reserve fund to cover a portion of the costs of penalties.** As explained earlier, no matter how well designed the performance measures are, random variation will create the potential that penalties could be triggered inappropriately. Currently, health insurance plans bear the full risk of higher spending on healthcare services, and they are expected and required to set premiums in excess of the amounts they actually spend on medical services in order to build and maintain reserves to manage this risk. If healthcare providers accept greater ac-

countability for spending, then they will also need to have similar reserves, and so the payments they receive for services under the APM could be increased by amounts that would enable the APM participants to build and maintain adequate reserves. The greater the potential penalties under the APM, the larger the reserve would need to be.¹⁶⁰

e. Using Bonuses in Addition to Penalties

Just as the impact of failure to achieve a Target will vary based on the amount by which utilization/spending falls short of the Target, the benefits of the APM will increase if APM participants not only achieve the Target, but perform even better. If a provider participating in an APM must pay penalties for failure to meet the Target, but receives no reward when performance is better than the Target, the participant will be penalized for random variation. Moreover, if extra time or cost is involved in achieving better performance, there will be little financial incentive for the APM participant to make that investment if its payment is the same no matter how much better its performance is.

This can be addressed by using bonus payments in addition to penalties, i.e., by making an additional payment to the provider if performance on the utilization/spending measure is better than the Target amount. For example, if the APM was expected to reduce the rate of avoidable hospitalizations by 10% but an APM participant actually achieved a 15% reduction, the APM participant could receive a bonus payment. As long as

the bonus is less than the additional savings the payer achieves as a result of the better performance, the payer would receive additional net savings and the provider would receive additional revenue.

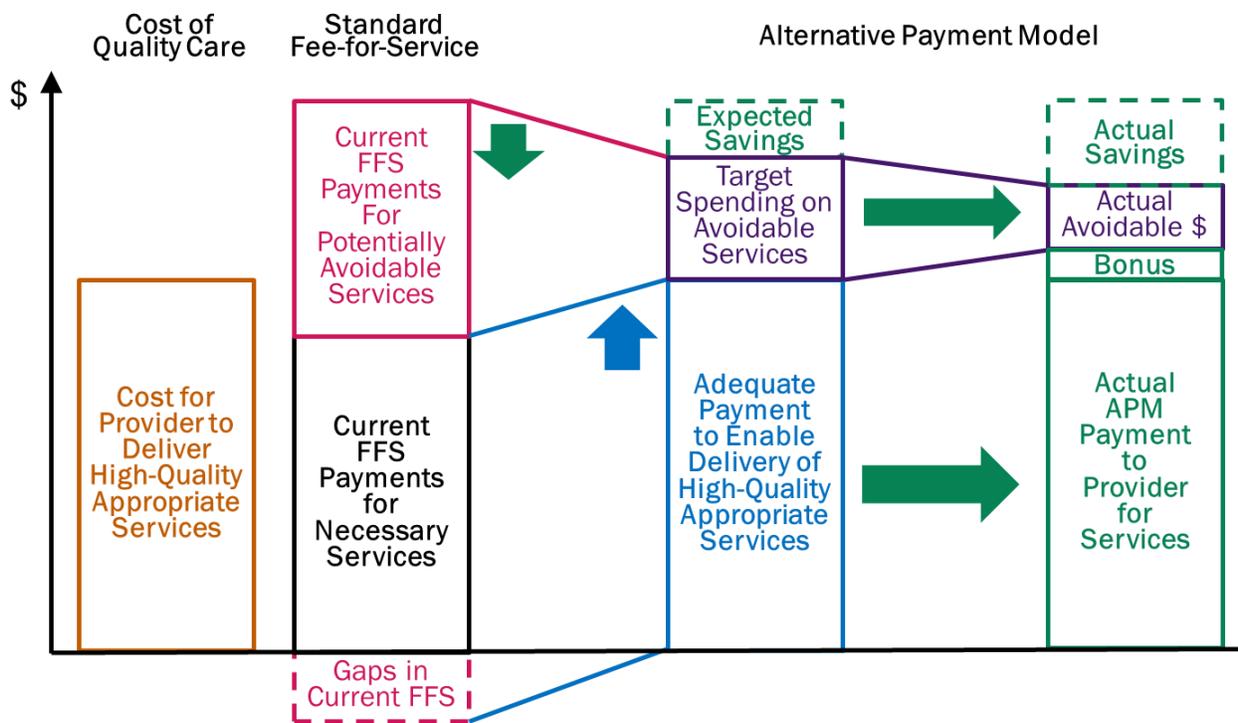
Bonuses vs. “Shared Savings”

Although paying a bonus when utilization/spending on a service is below a target amount could be described as “sharing savings” with the APM participant, this is very different than the payment model used in the Medicare Shared Savings Program and a number of other APMs. In the Medicare Shared Savings Program, there is no change in the amounts that providers are paid for delivering additional services in order to directly address the types of barriers in the payment system described in Section V. If an APM participant receives a shared savings payment, it is not a “bonus” on top of new or revised payments for desirable services, that is *the only payment* for such services. If the APM participant incurs costs to deliver new types of services which are not eligible for payment under existing payment systems, and the reduction in spending does not exceed the target savings amount, the payer saves money but the participant incurs a loss.

Symmetric vs. Asymmetric Bonuses and Penalties

The same types of alternative approaches described above for setting penalties can be used for setting bonuses. The methodology for bonuses and penalties

**FIGURE 7
BONUS (INCREASE IN COMPONENT #1 PAYMENT)
FOR PERFORMANCE BETTER THAN THE TARGET**



could either be *symmetric* or *asymmetric*. Under a symmetric bonus and penalty system, if performance is better than the Target by a certain amount, the bonus will be the same as if the performance was worse than the Target by the same amount. If the bonuses and penalties are asymmetric, then the penalty would be higher than the bonus or vice versa.

“Upside only” APMs are an extreme version of asymmetry, in which there is a bonus for performing better than the Target, but no penalty if performance falls short. For example, in the CMS Oncology Care Model, oncology practices in “Track 1” are paid a monthly payment and they receive a shared savings payment if spending is below the Benchmark, but they pay no penalty if spending is above the Benchmark.¹⁶¹

f. Adjusting Penalties Based on Performance on Multiple Measures

If two or more separate measures of utilization or spending are being used, a decision needs to be made as to whether penalties based on the measures will be determined separately or jointly. For example, if the APM participant fell short of the Target on one spending measure, but reduced spending by more than the Target on a different measure, penalizing the participant for poor performance on the first measure could be inappropriate if the net reduction in spending on the two measures combined was sufficient to achieve the overall spending goals of the APM.

It might seem that the solution to this is to simply combine the separate measures of spending into a composite measure and determine whether a penalty is warranted based on that single measure. However, this can be problematic for the same reasons that composite spending measures can be problematic; for example, if a random decrease in spending on a service over which the APM participant has little control happened to offset an increase in spending on a service that the APM participant directly controls, it could be inappropriate to impose no penalty simply because the combined spending on the two services showed savings.¹⁶² Instead, the APM participant’s performance would first need to be evaluated on each individual measure in order to determine whether performance had been better or worse than the Target for that specific measure. Then, if there was a negative deviation on one measure that would ordinarily trigger a penalty, but there was a positive deviation on another measure, an adjustment could be made to the penalty for the first measure based on the relative magnitude of the deviation on the second measure.

If bonuses are paid when performance is better than the Target as well as penalties for falling short of the Target, then it may be possible simply to net out the bonuses and penalties to determine the combined impact of performance on all of the measures, particularly if the bonuses and penalties are proportional to the amount of avoidable spending.

Example: Assume that the APM is supporting the ability of cardiology practices to deliver a new home care service for patients with heart failure in order to reduce avoidable ED visits and hospitalizations. Although there is a close connection between ED visits

and hospitalizations, it is desirable to evaluate performance on the two services separately, because the rate of hospitalizations depends not only on how often patients come to the ED but also on how severe the patient’s condition is and how likely the ED physicians are to admit the patient to the hospital, and the amount spent on the hospital stay depends not only on the severity of the patient’s heart failure, but also on any complications that occur due to the care delivered by the hospital staff. If the cardiology practice targeted its services on the patients at highest risk of hospitalization, it is possible that the practice would fail to reduce the overall rate of ED visits by the targeted amount, but it would reduce the rate of hospitalizations by more than the target amount. If the additional savings from the better-than-expected reduction in hospitalizations was greater than the shortfall in savings associated with the less-than-expected reduction in ED visits, then no penalty would be imposed on the cardiology practice. If both bonuses and penalties are paid for the changes in ED visits and hospitalizations and if the bonuses and penalties are proportional to the change in spending on ED visits and hospitalizations, then calculating the net bonus/penalty would achieve the same result.

Option 2: Outcome-Based Payment

As noted earlier, a significant weakness of Option 1 is that if a Population-Level Target is used to determine penalties, the payments for each patient will not be tied to the value of care that patient received. Not only is this undesirable, it means that Option 1 cannot easily be used for patients who are paying for their own care.

Even if a Patient-Level Target is used, if the penalty is less than the amount spent above the Target, the patient for whom the provider fails to achieve the Target could be paying more than the patient who does achieve the Target, which is the opposite of what a “value-based payment” is supposed to achieve.

Under Option 2, the APM participant would receive no payment of any kind under the APM for an individual patient unless (a) the provider delivered the planned services supported by the APM, and (b) the patient did not receive the planned or unplanned services that the APM was supposed to avoid. Continuing the example used in Option 1, the provider would only be paid for home nurse visits to a patient if the patient receiving the visits is not hospitalized. If the patient is hospitalized, the provider would not receive the payment for home nursing even if the patient received the service.

Option 2 is the equivalent of a full money-back guarantee on the service supported by the APM. If the Patient-Level Target for a patient is achieved, the payment for the service is made, and if the Target is not achieved, there is no payment for the service (or any payment that was made is refunded).

a. Adjusting the Payment for Desirable Services

Under Option 2, the amount the APM pays under Component #1 for planned services would need to be increased to reflect not only the cost of the services but the likeli-

EXAMPLE OF AN OUTCOME-BASED APM

In the example used earlier where a nurse is hired to make home visits to patients with a moderately severe chronic disease in order to reduce avoidable admissions to the hospital from 25% to 10%, rather than paying \$1,000 for the home visiting service for each patient regardless of whether they are hospitalized and imposing penalties or awarding bonuses based on the overall rate of hospitalization, the provider could be paid using an outcome-based payment instead. If the provider is only paid for those patients who are not admitted, then if only 90% of the patients are expected to avoid a hospital admission, the payments for those patients would need to cover the full cost of employing the nurse. If the provider is paid \$1,112 per year for each patient who receives the visits and does not have a hospital admission, the provider would cover its cost and the payer would still save money. The total payments would be \$100,080 (1,000 patients x 90% without admission x \$1,112), which is greater than the \$100,000 cost of the nurse, but less than the savings of \$150,000 from the avoided admissions.

The impact of this approach on a health insurance plan is the same as Option 1 (the Penalty approach) if the provider is successful in achieving the target improvement. If the provider achieves the target reduction of 15% in the hospital admission rate, under Option 1, the payer would pay \$100,000 for the nursing service (\$1,000 x 100 patients) and save \$150,000 on hospital admissions (100 patients x 15% avoided admissions x \$10,000), for net savings of \$50,000. Under Option 2 (the Outcome-Based approach), the payer would pay \$100,080 for the nursing service (\$1,112 x 90 patients not hospitalized) and save \$150,000 on hospital admissions, for net savings of \$49,920.

However, Option 2 has a very different impact with respect to an individual patient. Under Option 1, all of the patients pay \$1,000 per year (or have that amount paid on their behalf by their health insurance plan), but the 10% of patients who are still hospitalized also have to pay \$10,000 for the hospitalization, so they spend 10% more than they would have before for the same outcome. Under Option 2, 90% of the patients would pay \$1,112 per year, and the 10% of patients who are hospitalized would pay only \$10,000; they would be refunded for the home visits that failed to prevent their hospitalization. The difference is similar if the patient is paying cost-sharing rather than the full payment for services. For example, in the case of Medicare beneficiaries who would pay 20% cost-sharing for the home care service and a \$1,340 deductible for an inpatient hospitalization, the beneficiary who was hospitalized would pay \$1,540 under Option 1 but only \$1,340 under Option 2, a \$200 savings. The beneficiary who was not hospitalized would pay \$200 under Option 1 and \$222 under Option 2, a \$22 difference. Compared to the fee-for-service system, however, the patient would have a much lower likelihood of being hospitalized.

	FFS	Outcome-Based APM		
	Current Performance	Target Performance	Failure to Meet Target	Better Than Target
Patients				
Total # of Patients	100	100	100	100
% Patients Hospitalized	25%	10%	15%	5%
# Patients Avoiding Hospitalization	75	90	85	95
Payment Per Patient for Nurse Home Visits		\$1,112	\$1,112	\$1,112
# Patients Hospitalized	25	10	15	5
Payment Per Patient for Nurse Home Visits		\$0	\$0	\$0
Provider Revenue/Cost				
Revenue from Payments for Nurse Home Visits		\$100,080	\$94,520	\$105,640
Cost of Nurse Making Home Visits		(\$100,000)	(\$100,000)	(\$100,000)
Profit/Loss for Provider for Nurse Home Visits		\$0	(\$5,480)	\$5,640
Payer Spending				
Spending on Hospitalizations	\$250,000	\$100,000	\$150,000	\$50,000
Payments for Nurse Home Visits		\$100,080	\$94,520	\$105,640
Total Payer Spending	\$250,000	\$200,080	\$244,520	\$155,640
% Change in Spending		-20%	-2%	-38%

hood that the APM participant will achieve the Target. If the payment amounts for the desirable services under Component #1 were based solely on the cost of delivering those services to the patients who receive them, then the total payments made for the patients who do achieve the result will fall short of the total cost of delivering the services unless 100% of the patients achieved the desired result. Instead, the Component #1 payment amount needs to be defined by determining the total cost of delivering the service to all of the patients who are likely to receive it and dividing by the number of patients *who would be expected to achieve the result*.

Because of the need for this kind of adjustment, the Outcome-Based Payment approach is most appropriate where the Patient-Level Target can be achieved in very high percentage of cases, i.e., where the equivalent Population-Level Target for an undesirable service would be close to zero or where the probability of random or uncontrollable deviations from a total spending level would be very low.

As discussed in detail in Sections VI.B.1 and VI.B.2, this is more likely to be possible when utilization/spending measures are focused on types of services and spending the APM provider can control and when the measures are stratified to reflect significant differences in patient needs and risks.

b. Penalties and Bonuses Under Outcome-Based Payments

Making payments contingent on performance automatically results in penalties for poor performance and bonuses for good performance similar to what would be achieved using Approach #1 under Option 1. If performance falls short of the Target that was used to establish the Component #1 payment amount for the desirable services, then the provider will receive less revenue, the same as if they had paid a penalty proportional to the payments they receive for their services. If performance is better than the Target, the provider will receive more revenue, the same as if they received a bonus proportional to their payments.

The default approach would be no payment for an individual patient if the target result was not achieved, which is equivalent to a penalty that is 100% of the payment. However, it is also possible to define a partial money-back guarantee, i.e., to say that the provider would refund a fraction of the payment for the desirable services if the target result is not achieved.

c. Sharing Responsibility for Performance by the Provider and Patient

In many cases, the desired performance is not completely in the provider's control. For example, having a nurse make a home visit to a patient with chronic disease to provide education about their medications and how to avoid exacerbations can help the patient avoid hospitalizations, but success also depends on the patient taking their prescribed medications properly and following the nurse's advice about how to avoid exacerbations. However, when the desired result isn't achieved, it isn't easy to determine whether the fault rests with the provider or the patient. In the example, the patient might not have

taken their medication because (1) the nurse failed to explain the importance of doing so, (2) the nurse or physician failed to recognize the side effects the patient was experiencing and failed to make appropriate adjustments in their medications, or (3) the patient simply wasn't willing to take their medications.

Because Option 2 defines the payment adjustments at the level of the individual patient, it provides a way for the provider and the patient to share responsibility for performance when a successful result depends on actions by both of them. For example, if the provider delivering home services is paid 50% of the standard payment when a patient is hospitalized, then both the patient and the provider will take part of the financial risk for the success of the home visits. If a hospitalization occurs, the provider will only be paid for a portion of the service, reflecting their partial responsibility for the poor outcome.

d. Limiting the Provider's Losses

Under Option 1, it is relatively easy to limit the total financial risk for an APM participant, because the APM participant pays a single aggregate penalty to a payer at the end of a performance period, and that penalty can be limited to a maximum amount. Under Option 2, however, the penalty is determined on a patient-by-patient basis, and the size of the gap between total payments and total costs will change each time that a new patient is treated.

Consequently, limiting the total financial risk for the provider under Option 2 requires use of a stop-loss mechanism. Under stop-loss insurance, if the APM participant reaches a certain threshold of losses, the participant receives additional funds from an insurer to cover all or part of the losses above that threshold. If the stop-loss protection is not provided by the patient's health insurance plan as part of the APM, the APM participant would need to purchase stop-loss insurance from an insurance company, and the premium that the APM participant pays for the stop-loss insurance would then need to be factored into the cost of delivering services under the APM.¹⁶³

Example: In the example earlier, 25% of patients with a chronic disease are currently admitted to the hospital during the course of the year, healthcare providers hire nurses who each provide home visits for 100 patients, the home visits are expected to reduce the hospitalization rate to 10%, and the providers are paid \$1,112 per year for each patient who does not have a hospitalization. If the home visits achieve the expected results, the provider would be paid \$100,080 per year for 100 patients (100 x 90% x \$1,112), which would be sufficient to pay for the nurse's compensation, travel, etc. However, if the hospitalization rate for the patients is higher than 10%, the provider's revenues would fall short of the cost of the nurse. If providers were unwilling to risk a loss of more than \$15,000, the payer could provide an additional payment at the end of the year if the hospitalization rate was higher than 23% (if the hospitalization rate was 24%, then the provider would not be paid for 24% of patients instead of

10% of patients, causing a loss of $14\% \times 100 \times \$1,112 = \$15,568$). If the provider needed to purchase stop loss insurance from a separate insurer, that insurer would determine the premium per patient and the \$1,112 payment to the provider would need to be increased by the premium amount.

Option 3: Bundled/Warrantied Payments

Option 2 assures each individual patient and their health insurer that if the planned services supported by the APM fail to achieve the desired impact on utilization and spending, the patient/payer will not have to pay for the planned services that are supported by the APM. However, the patient and/or payer would still have to pay for the unplanned services or increases in spending they had expected to avoid. If the planned services had a negative impact on the patient (e.g., the patient experienced an infection or complication), the patient and/or payer would have to spend more money for other services in order to address the new problems.

Approach #2 in Option 1 addresses this for a payer by requiring the APM participant to pay a penalty based on the amount of spending that was not avoided and/or an increase in spending on related services beyond what would otherwise have been expected. However, this population-based penalty approach does not translate into an assurance that the spending on avoidable services for each individual patient will be lower than it would otherwise have been.

Option 3 would address this by defining a bundled/warrantied payment for each patient that supports the ability of the APM participant to deliver the planned services but also requires the APM participant to pay for the unplanned services if they are actually needed and delivered. Continuing the example used earlier in this section, if an APM is intended to support home care services for patients with chronic disease in order to reduce avoidable hospitalizations, the APM participant would use the bundled/warrantied payment to support the home care service but it would also need to pay for any avoidable hospitalizations that the patient experienced using the funds from that same payment.

a. Bundling for Accountability vs. Flexibility

Bundled payments were previously discussed in Section VI.A as an optional approach for paying for new or revised services. In that case, the primary purpose of bundling was either to (1) provide *flexibility* for the APM participant to deliver different combinations of desirable services without requiring a justification for the specific number or types of services that were delivered and without any concern that a particular combination would have a negative impact on the revenues needed to support the services, or (2) control the services delivered by another provider. All of the services included in such a bundled payment would ordinarily be planned services that the APM participant would deliver itself or would order from other providers and pay for through the bundled payment.

Under Option 3, the primary purpose of bundling is to create *accountability* for the APM participant. Here, the

bundle would be used to pay for the planned services and also the unplanned services if they occur. Since the unplanned services might be delivered by providers other than those participating in the APM, the bundled/warrantied payment under Option 3 would need to be a multi-provider bundled payment (Option 12 in Section VI.A) even if the planned services were being delivered by a single provider.

Option 3 is analogous to a warranty on a product or service. The APM participant is not guaranteeing that no complications will occur, it is merely agreeing to pay to treat them if they do occur without receiving any additional payments from the patient or payer. This is similar to a product warranty; for example, the warranty on a new car does not guarantee that the transmission won't fail, it merely assures the car buyer that if the transmission fails, it will be repaired at no additional cost to the car buyer.

In addition, most product warranties are *limited* warranties, i.e., there are certain circumstances in which repairs will involve an extra charge (e.g., repairs needed due to the owner's failure to carry out required product maintenance). In the case of the APM, these decisions about what is included and excluded are made in defining the spending measures and were discussed in Section VI.B.1.

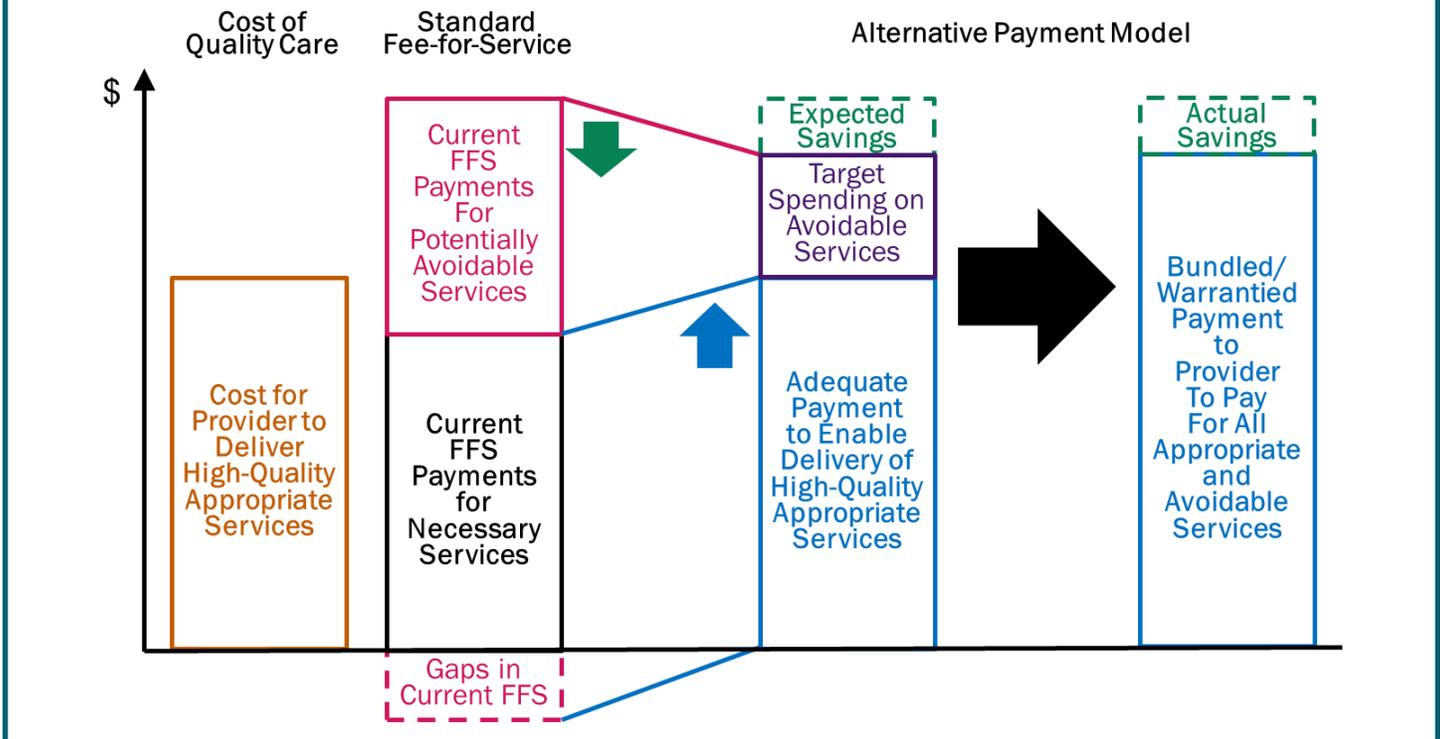
Example: The Geisinger Health System's ProvenCareSM program has defined bundled/warrantied payments for several procedures, beginning with coronary artery bypass graft surgery in 2006. The payment covers pre-operative evaluation and workup, all hospital and physician fees, all routine post-discharge care, and management of all related complications, including hospital readmissions for post-operative complications that occur within 90 days.¹⁶⁴

Example: In 1987, an orthopedic surgeon in Lansing, Michigan and the local hospital, Ingham Medical Center, offered a fixed total price for surgical services for shoulder and knee problems and a warranty for any subsequent services needed for a two-year period, including repeat visits, imaging, rehospitalization and additional surgery. The results were that the health insurer paid 40% less than otherwise, the surgeon received over 80% more in payment than otherwise, and the hospital received 13% more than otherwise, despite fewer rehospitalizations.¹⁶⁵

b. Determining the Amount of the Bundled/Warrantied Payment

Since the APM participant would have to use the bundled/warrantied payment to pay for both the planned services and the unplanned services if the patient needs them, the amount of payment under Option 3 will need to be larger than the amounts under Options 1 and 2 where the payment only needed to cover the costs of the planned services. The performance Targets provide a mechanism for determining how much higher the Option 3 payment needs to be.

**FIGURE 8
BUNDLED/WARRANTIED PAYMENT**



c. Limiting the Provider's Losses

Under Option 2, if performance is poor, the largest amount the APM participant could lose is the cost the participant incurred for delivering the planned services. However, under Option 3, the loss for a poorly-performing participant could be dramatically higher. In the example above, if the APM participant was delivering home visits to 100 patients, and 100% of the patients had a hospitalization (rather than the 5% target rate or the 20% baseline rate), the participant would be paid \$150,000 (100 patients times the \$1,500 payment per patient), but would incur costs of \$1.1 million (the \$100,000 cost of the nurse plus a \$10,000 payment for each of the 100 hospitalizations), resulting in a loss of nearly \$1 million. In fact, some or all of the patients could potentially be hospitalized multiple times, increasing the total losses even more.

There are two different mechanisms that can be used to limit the total financial risk for the provider under Option 3:

- **Outlier payments.** The APM Participant could receive an additional payment if an individual patient needed an unusually large number of unplanned services or unusually expensive services. This would be similar to Option 9 that is discussed in Section VI.A. Alternatively, the APM Participant could purchase “individual stop-loss” insurance and treat the premium for that insurance as part of the cost of delivering the planned services.

- **Risk Corridors/Aggregate Stop-Loss Insurance.** Even if no individual patient requires a large number of unplanned services, the APM participant can experience financial problems if unplanned services are needed by a large number of patients. The APM participant could purchase “aggregate stop loss” insurance to protect against this, or the payer could do the equivalent by defining what is called a “risk corridor,” i.e., if the total amount the APM participant has to pay for unplanned services exceeds a certain threshold, the payer would make an additional payment.

d. Implications for Insurance Regulation

Under Option 3, the provider could potentially be viewed as offering “insurance” by state insurance commissioners. This would be particularly true if the warranty covers services that the APM provider does not deliver directly, since that could involve a large out-of-pocket expense for the provider, but even if the warranty simply covers services the APM provider delivers, the provider might not be able to fulfill the warranty if a large number of patients experience the complications or other situations that are covered by the warranty. The lower the threshold for stop-loss insurance, the less insurance risk that the APM provider is taking on, but the APM provider would then need to pay a higher premium for the associated stop-loss insurance and that would reduce the opportunity for the provider and the primary payer to achieve savings.

EXAMPLE OF A BUNDLED/WARRANTIED APM

In the example used earlier, where a nurse is hired to make home visits to chronic disease patients in order to reduce avoidable admissions to the hospital, the average cost per patient of delivering the service would be \$1,000. As shown in the table, if delivery of the service is expected to reduce the rate of hospitalizations to 10%, and if a hospital admission costs \$10,000, then the payment per patient would have to be increased by \$1,000 (\$10,000 x 10%) to \$2,000 in order for the APM participant to have enough revenue to pay for the hospitalizations that do occur as well as the home nursing service.

Although this makes the home care service twice as expensive as under Option 1 for the patient who would not have been hospitalized, it reduces the total spending for the patient who would have been hospitalized by 82% (\$2,000 compared to \$1,000 + \$10,000). Every patient (and their payer) would have the certainty of paying \$2,000 for both home visits and hospitalizations, rather than either \$1,000 or \$11,000. Moreover, if the rate of hospitalizations without the home visits was 25%, then on average, a patient and their payer would have expected to spend \$2,500 (\$10,000 for a hospitalization times a 25% chance of being hospitalized), so the \$2,000 payment is 20% less. Overall, spending is \$500 lower per patient than it would otherwise have been, which generates the same \$50,000 savings for a payer with 100 patients as under Option 1, but Option 3 can also be used for patients who are paying directly (without insurance) and to control the cost-sharing for patients who do have insurance.

From the provider's perspective, the penalty for failure to achieve the Target is much higher, because the provider is responsible for paying for each \$10,000 hospitalization, rather than simply losing part or all of the payment for the nurse home visits. On the other hand, the reward for performing better than the Target is much larger.

	FFS	Bundled/Warrantied APM		
	Current Performance	Target Performance	Failure to Meet Target	Better Than Target
Patients				
Total # of Patients	100	100	100	100
% Patients Hospitalized	25%	10%	15%	5%
# Patients Avoiding Hospitalization	75	90	85	95
Bundled/Warrantied Payment		\$2,000	\$2,000	\$2,000
# Patients Hospitalized	25	10	15	5
Bundled/Warrantied Payment		\$2,000	\$2,000	\$2,000
Provider Revenue/Cost				
Revenue from Bundled Payments		\$200,000	\$200,000	\$200,000
Cost of Nurse Making Home Visits		(\$100,000)	(\$100,000)	(\$100,000)
Payments for Hospitalizations		(\$100,000)	(\$200,000)	(\$50,000)
Profit/Loss for Provider		\$0	(\$50,000)	\$50,000
Payer Spending				
Spending on Hospitalizations	\$250,000			
Spending on Bundled/Warrantied Payments		\$200,000	\$200,000	\$200,000
Total Payer Spending	\$250,000	\$200,000	\$200,000	\$200,000
% Change in Spending		-20%	-20%	-20%

Option 4: Termination of the Provider's Participation in the APM

Options 1-3 all assume that a provider participating in an APM that fails to meet the Target for one or more performance measures will pay some type of financial penalty and continue participating in the APM (if they wish to do so). A fourth option is to simply terminate the provider's participation in the APM altogether if that provider does not achieve success on the performance measures.

This is essentially the approach that currently exists in the standard fee-for-service payment systems used by Medicare and private payers. A provider must meet some set of minimum standards of performance (e.g., hospitals must be accredited) in order to continue being paid for services. For many aspects of performance that are assessed through this process, there is no absolute threshold of performance that must be achieved, but instead a determination is made as to whether the problems are both remediable and being remedied in order to decide whether termination should occur. A variation on this approach could be used in APMs.

Example: In the Independence at Home Demonstration, Congress required that a physician practice be terminated from participating in the program if "(A) the Secretary estimates or determines that such practice will not receive an incentive payment for the second of 2 consecutive years under the demonstration program; or (B) such practice fails to meet quality standards during any year of the demonstration program."¹⁶⁶

Example: In Track 1 of the CMMI Oncology Care Model (OCM), there is no financial penalty for the participating practice if total spending on the patients is higher than the target spending level defined in the payment methodology. However, the program requirements state that if an oncology practice does not achieve the goal of a 4% reduction in spending by the third year of participation, the practice will not be permitted to continue participating in the OCM alternative payment model.¹⁶⁷

From the payer's perspective, Option 4 precludes any further financial losses resulting from the provider's shortfall in performance. However, it also precludes any opportunity to recoup the losses that have occurred so far or to benefit from any future savings the APM participant might have generated if performance improved.

From the APM participant's perspective, there is still a financial penalty associated with termination of participation if the entity had hired new staff, purchased new equipment, or incurred other kinds of costs that were not fully covered by the payments received prior to termination.

An advantage of Option 4 is that it allows greater flexibility to consider the circumstances that may have led to failure or success in meeting the targets. For example, if an APM participant's patients required a higher-than-average number of expensive services due to characteristics for which there was not effective adjustment in establishing the Benchmark and Target, a determination

could be made that no penalty is justified, since the payer would have spent a higher amount for the care of the patients in the absence of the APM. Conversely, if it turns out that the Target spending level was achieved by systematically avoiding patients who were likely to require multiple services, the APM participant could be terminated even though a standard formula might have determined that a bonus should be paid.

Option 4 may be the only feasible option for small providers, for APMs focused on small numbers of patients or health conditions that occur relatively rarely, or for APMs designed to reduce problematic outcomes that occur rarely or unpredictably. No matter how sophisticated the statistical methodology, it may be impossible to create a fair way of determining penalties or basing payments on outcomes, and so a different form of evaluation may be necessary. For example, outcomes might be evaluated over a multi-year period, or detailed clinical audits of individual patients might be used to verify that the most appropriate care was being delivered.

Option 4 could also be used during the initial years of implementation of an APM, or during the type of "beta testing" process described in Section VIII.B, when there is uncertainty about how quickly changes in care delivery can be implemented and how much of an impact those changes will have on spending or quality. A transition could then be made to one of the other Options.

Option 5: Termination or Modification of the APM

A final option is to stop using the APM altogether. If APM participants collectively are not succeeding in reducing spending or maintaining spending while improving quality, then it makes sense to modify the design of the APM or to terminate it and develop something different.

The statute creating the Center for Medicare and Medicaid Innovation (CMMI) requires that CMS "terminate or modify the design and implementation of a model unless the Secretary determines (and the Chief Actuary of the Centers for Medicare & Medicaid Services, with respect to program spending under the applicable title, certifies), after testing has begun, that the model is expected to (i) improve the quality of care (as determined by the Administrator of the Centers for Medicare & Medicaid Services) without increasing spending under the applicable title; (ii) reduce spending under the applicable title without reducing the quality of care; or (iii) improve the quality of care and reduce spending. Such termination may occur at any time after such testing has begun and before completion of the testing."¹⁶⁸

For example, CMMI implemented the Comprehensive Primary Care Initiative (CPCI) APM for primary care practices in 2012. The participating practices received higher payments to deliver additional services, but the APM did not make each practice directly accountable for reducing spending by at least as much as the increased payment. An evaluation conducted by CMMI concluded that CPCI had increased overall Medicare spending¹⁶⁹, so the CPCI APM was terminated at the end of 2016 and replaced by the Comprehensive Primary Care Plus APM, which differs from CPCI in significant ways.

**TABLE 10
METHODS OF PERFORMANCE-BASED PAYMENT FOR UTILIZATION/SPENDING**

Method	Strengths	Weaknesses
1a. Penalties Based on Failure to Meet Targets	The amount of the penalty can be made small or large based on factors such as the likelihood of achieving the Target and the ability of the provider to afford penalties	<p>The provider could be at significant financial risk if the size of the penalty is based on the amount of spending in excess of the Target and if the Target spending is much larger than the provider's revenue</p> <p>A payer's net spending could increase if the penalty is less than the increase in spending beyond the Target</p> <p>Some patients and payers will have to pay more for the services they receive even if a Population-Level Target is achieved</p> <p>There is no reward for performance that is better than the Target</p>
1b. Penalties & Bonuses Based on Achieving Targets	Encourages generating more savings than the Target Change	Bonuses for high-performing APM providers could potentially exceed the amount of savings and penalties from low-performing providers, thereby increasing total spending
2. Outcome-Based Payment	<p>Ensures there is no payment (or a large reduction in payment) for an individual patient if the Target is not achieved for that patient</p> <p>Results in the equivalent of a bonus for high performance and a penalty for poor performance</p>	<p>Works best when there is a high probability of achieving the Target for most patients</p> <p>Does not compensate the patient or payers for costs they incur for treating complications and other costs not directly supported by the payment</p>
3. Bundled/Warrantied Payment	Can protect the patient and payer from spending more in total since spending on both intended and unintended changes in services are included	Could create significant financial risk for providers if the unplanned services covered by the warranty are expensive or occur frequently
4. Termination of Provider's Participation in the APM	<p>Allows the determination of whether savings were achieved to be based on a more detailed evaluation than is possible through a comparison of actual spending to the Target</p> <p>Enables customized evaluations to be performed for small providers where utilization or spending measures are unreliable</p>	<p>May not be possible to accurately determine the reasons for high utilization or spending</p> <p>Could result in higher spending before a decision is made to terminate a provider</p> <p>Premature termination could result in a lost opportunity to achieve savings in the future</p>
5. Termination/Modification of the APM	Avoids continued losses from an APM that fails to achieve its goals	Providers could be less likely to participate in the APM or to transform care delivery significantly if it is uncertain whether the APM will continue to be offered in the future or if there is concern that major changes will be made

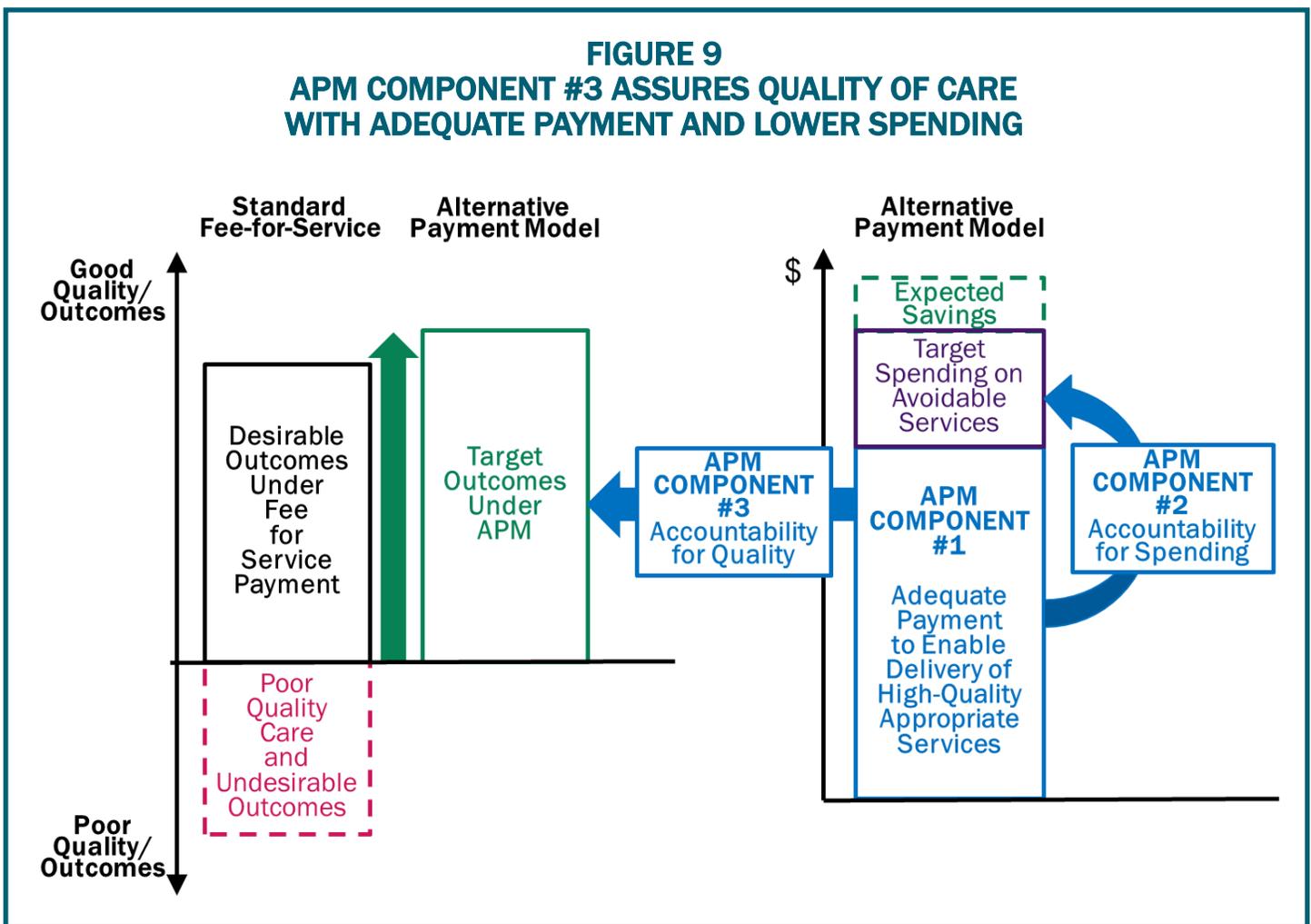
C. APM Component #3: Creating Accountability for Quality

As explained in Section II, in order to qualify as an Alternative Payment Model, it is not enough to simply maintain or reduce spending; the quality of care for patients must also be maintained or improved. Consequently, in addition to Component #1, which is designed to address the types of payment barriers described in Section VI.B, and in addition to Component #2, which is designed to ensure that spending is maintained or reduced, an Alternative Payment Model must include a third component for ensuring that quality does not worsen. If the APM is explicitly intended to improve quality without increasing spending, Component #3 must provide a way of ensuring that improvement occurs.

An accountability component for quality in an APM has four distinct elements, similar to those for Component #2:

1. One or more *measures of quality* that need to be maintained or improved by the services supported by the APM;
2. *Targets* for the level of quality that must be maintained or the improvement that must be achieved on each aspect of quality in order for the APM to be deemed successful in achieving its goal;
3. A *performance assessment methodology* to determine whether a specific entity participating in the APM has achieved the Quality Targets; and
4. A *mechanism for adjusting payments based on performance*, i.e., what changes will be made in payments if the Targets are not achieved.

FIGURE 9
APM COMPONENT #3 ASSURES QUALITY OF CARE WITH ADEQUATE PAYMENT AND LOWER SPENDING



1. Defining the Accountability Measures

Designing a quality accountability component for an APM is challenging because “quality” has many dimensions. Different patients may benefit in different ways from new approaches to care delivery. There are also multiple ways that patients could be harmed when different services are being delivered and when physicians or other healthcare providers are subject to financial penalties based on utilization and spending.

Many current APMs have chosen to hold APM participants accountable only for aspects of quality where quality measures already exist and are being reported, even if those measures do not match the specific aspects of quality likely to be affected by the APM. Although there is no requirement that an APM use only existing quality measures¹⁷⁰, use of existing measures has been promoted as a way to reduce the burden and expense of collecting additional data and reporting more or different measures.

However, holding providers accountable for quality measures that are not directly aligned with the goals of the APM actually increases the burden on the APM participants. Not only must they deliver the services the APM is designed to support and ensure that the APM’s goals for reducing utilization and spending are met, they may need to undertake additional, unrelated activities in order to improve performance on unrelated quality measures. Moreover, the measures may not protect patients against the kinds of problems in quality that the APM is likely to cause.

In some cases, there will be existing quality measures that match the needs of the APM; in other cases, one or more new quality measures will need to be developed. Determining whether existing or new measures are most appropriate requires three separate steps:

- a. Identifying the aspects of quality that could be affected by the APM;
- b. Determining how changes in those aspects of quality should be assessed; and
- c. Determining whether and how data needed to make such assessments can be obtained.

a. Identifying the Aspects of Quality Where Accountability is Needed

There are four general areas that should be examined to determine what quality measures are needed for an APM:

- Aspects of quality where the APM is intended to make improvements;
- Aspects of quality that could be harmed by changes in services supported by the APM;
- Aspects of quality that could be harmed by the payment methodology or spending accountability components of the APM; and
- Aspects of quality necessary to ensure accurate payment under the APM.

i. Achieving Intended Improvements in Quality

If the APM is explicitly intended to improve one or more aspects of the quality of care in addition to or instead of reducing spending, then those specific aspects of quality need to be clearly defined and accurately measured so a determination can be made as to whether they are, in fact, being improved.

ii. Maintaining Quality of Care Achieved Under the Current Payment System

On the other hand, if the goal of the APM is to reduce spending, then the APM is required to *maintain* quality, but not necessarily to *improve* quality. In this case, multiple measures may be needed to ensure that current levels of quality are not being harmed. Even if the APM is focused on reducing utilization of a service that is either unnecessary or undesirable for *some* patients, that service may have benefits for *other* patients and it will be important to ensure that those patients are not harmed. For example, reducing the rate of hospitalizations has the potential to improve outcomes by reducing the number of patients exposed to risks of infection, falls, and medical errors that can occur in the hospital. However, reducing the rate of hospitalizations also has the potential to increase the number of patients who die or experience other problems that could have been prevented by a hospitalization. Even if it is viewed as an acceptable tradeoff to have a small increase in the rate of some adverse outcomes in order to achieve important benefits from reducing unnecessary or avoidable services, it would be important to know whether the increase in adverse outcomes is within the range that was expected.

Determining whether an APM preserves current levels of quality requires defining (and then measuring) the benefits that patients are receiving from the services delivered under the current payment system. For many services, there are not even clear definitions of these benefits, much less ways of measuring them, which makes it hard to determine whether quality is being improved or worsened under an APM.

In many cases, an APM will reduce the use of one service by enabling use of one or more new services or by encouraging greater use of one or more existing services. However, the new or expanded services delivered under the APM may have the potential to cause new or different problems for patients. For example, if a lower-cost knee implant is used for knee replacement surgery, but the patients receiving the implant experience new types of pain or limitations on mobility, the quality of care will have decreased even though overall savings have been achieved.

When new services are being created under an APM, or when existing services are being expanded to new types of patients, it may not always be possible to identify all of the potential quality problems in advance. Consequently, mechanisms may need to be created to identify what types of problems occur before consideration can be given to creating measures of the frequency of those

problems. For example, a process may be needed for contacting patients or family members to identify whether adverse events have occurred.

iii. Potential Quality Problems Created by the Payment Methodology

Although fee-for-service payment is criticized because healthcare providers are paid more when they deliver more services, this is actually a strength for patients who need more services than average. Using bundled payments rather than traditional fee-for-service payments can benefit patients by giving providers more flexibility about the ways in which they deliver services, but bundled payments also create the potential for patients to be underserved. For example, paying a primary care practice a monthly payment for each patient rather than a payment tied to office visits gives the practice the flexibility to deliver different kinds of services to patients. However, the monthly payment also means that the practice would be paid the same amount whether or not it scheduled an office visit that a patient needed, and the practice could try to avoid enrolling patients who require a large number of visits.

To address this, the APM should have a way of determining whether patients are receiving all of the services they need, particularly if a provider participating in the APM is no longer paid more for each service that is delivered.

In addition, the quality accountability component of an APM cannot be designed independently of the spending accountability component because the spending accountability component itself has the potential to harm the quality of care. The broader the measure of spending or utilization that a provider is accountable for controlling or reducing, the more ways there are for the provider to reduce spending other than by changing the specific kinds of services that the APM was intended to change, and the negative impacts of those other types of spending reductions must also be considered.

For example, if the goal of the APM is to reduce avoidable hospitalizations for exacerbations of chronic disease, but the APM holds participating providers accountable for the *total* rate of hospitalizations (not just the rate of *avoidable* hospitalizations associated with exacerbations), then one APM participant might achieve savings by reducing admissions for chronic disease exacerbations, but another APM participant might achieve savings by reducing hospital admissions for high-value elective procedures. The patients of the first provider might benefit from the change, but the patients of the second might not.

In theory, one could avoid the potential for undesirable reductions in services in other areas by measuring the quality of care in those areas and requiring that quality be maintained or improved. However, since quality measures are specific to the type of condition being

managed and the type of service being delivered, the more ways there are to reduce a measure of spending, the more measures of quality will be needed. Moreover, since there is also a cost to measuring quality, the costs and administrative burdens of the APM will increase if the APM tries to hold providers accountable for a broad measure of spending while protecting patients from adverse impacts.

Problems with Measuring Quality in Total Cost of Care Models

The most extreme version of this problem occurs with APMs that include a “total cost of care” accountability component. Requiring a decrease in the total amount spent on patients’ care in order to declare that the APM is a success may give the payer greater assurance that a reduction in one type of avoidable spending is not being offset in other ways, but there is also the risk that it could lead to stinting on quality in any of a wide variety of ways. In order to assure each individual patient that the provider’s accountability for total spending was not harming them, a very large number of quality measures could be needed.

Current APMs that require participants to reduce the total cost of care typically only attempt to measure a subset of the potential ways in which the quality of care for patients might be harmed. The implicit assumption is that if a healthcare provider performs well on a representative subset of quality measures that are explicitly monitored in the APM, the provider is likely to

also deliver high-quality care on all other measures. However, even if there is a high correlation in a provider’s performance on multiple quality measures under the *current* payment system, there is no assurance that the correlation will remain similarly high under the APM. If the participants in an APM are expected to reduce total spending on their patients, and if they are held accountable for maintaining or improving the level of quality for health condition A but not for health condition B, then some APM participants might concentrate their efforts to reduce spending on patients who have condition B, and any negative impacts on quality there would go unnoticed and unpenalized. The correlation in quality between condition A and condition B that existed before implementation of the APM would then disappear as a result of the APM.

Example: In the Medicare Shared Savings program, 31 different measures are used to assess the quality of care delivered to patients, but there are no measures of the quality of cancer treatment, despite the fact that cancer treatment can be very expensive and some effective cancer treatments can have serious side effects and additional spending is required to treat the side effects. Using less expensive therapies with fewer side effects to treat cancer could significantly reduce spending in an ACO but could also harm patient outcomes, and there is no direct

*way to measure or prevent that with the measures that are currently being used.*¹⁷¹

This problem can be mitigated or avoided if the spending accountability component of the APM uses measures specifically focused on the types of spending that the APM is expected to reduce. If the APM participant only receives “credit” for reducing spending on specific types of services or health conditions, then the primary place where quality measures are needed are for aspects of quality that could be negatively affected by changes in care in those specific areas.

This is also a reason for avoiding excessively stringent standards of statistical significance in evaluating whether spending in other areas has changed. For example, if the APM requires that there be no significant increase in spending on unrelated services, an APM participant (particularly a participant with a small number of patients) may feel compelled to find ways of reducing spending on the unrelated services simply to ensure that uncontrollable and random factors do not make it appear that unrelated spending has increased. This then could require the use of additional quality measures to ensure that those spending reductions do not cause reductions in the quality of care for patients.

iv. Ensuring Accurate Payment Under the APM

In addition to issues associated with the quality of care each patient receives and the outcomes they achieve, it will also be important to ensure the accuracy of the information used to determine the correct payment and to measure performance as part of the APM. For example:

- In a condition-based payment for a particular disease or combination of diseases, it will be important to ensure the accuracy of the diagnosis on which the payment is based. Currently, most payments are based on the type of service delivered, but if payments are based on diagnoses, processes will be needed to ensure the accuracy of the diagnoses.
- If a payment model requires accountability for maintaining or improving a particular outcome for a patient, it will be important to ensure that the process used to measure that outcome is accurate and reliable.

Assessing the accuracy of diagnoses and outcome information can also be considered “quality measures,” since inaccuracies will have negative effects on patients as well as resulting in errors in payments under the APM.

b. Determining How to Assess a Particular Aspect of Quality

i. The Benefits and Challenges of Outcome Measures

There is widespread agreement that it would be preferable to measure the quality of health care based on the outcomes achieved for patients rather than to simply measure which services a patient received and/or how those services were delivered.

In MACRA, Congress expressed a preference for using outcome measures.¹⁷² However, relatively few such measures have been developed and even fewer are currently in use. There are several reasons why outcome measures are not more widely used:

- They can be expensive to collect, particularly if they require finding and surveying patients who are no longer actively receiving care;
- If the outcomes of interest occur long after services are delivered, there will be a significant delay in determining whether outcomes have improved or worsened;
- Serious negative outcomes (such as death) occur relatively infrequently, so changes in these outcomes are difficult to measure accurately, particularly with small numbers of patients; and
- Many outcomes are affected as much or more by factors unrelated to healthcare services as they are by anything that a healthcare provider can do, so it is difficult to determine whether changes in outcomes were or were not due to actions by a healthcare provider.

APMs could potentially facilitate the collection of more outcome measures by providing a means and a rationale for paying for the cost of data collection. For example, in the Comprehensive Care for Joint Replacement APM, the quality scores for participating hospitals are increased by 10% if they submit data on patient-reported outcomes and patient risk factors. The risk factors represent information such as body mass index, pre-operative use of narcotics, and level of pain, and the outcomes measured include pain and functionality.¹⁷³

ii. Intermediate Outcomes as a Proxy for Outcomes

A commonly used alternative to measuring outcomes is to measure the results of laboratory tests and other biomarkers. These are often referred to as “intermediate outcome” measures, since they reflect some kind of a change in the patient, but not the ultimate change that is truly the goal of care. If evidence shows the intermediate outcomes are highly correlated with the longer-term outcomes of interest, it may be more practical to evaluate performance based on the intermediate outcomes than on the actual outcomes. However, if the intermediate outcomes are not directly related to the outcomes that are truly desired, holding providers accountable for achieving them could harm patients and divert time and effort away from delivering services that would have a bigger impact on spending and quality.

iii. Processes as a Proxy for Outcomes

The most commonly used quality measures continue to be “process” measures, i.e., measures of whether a particular activity was performed, such as administering a particular test or drug. As with intermediate outcome measures, process measures can be a desirable alternative to outcome measures if (but only if) there is strong evidence showing that achieving the process measures is closely correlated with the desired outcome. For purposes of accountability in an APM, it will

generally be far easier to measure whether a provider delivered desirable services in a specific way than to measure what outcome was achieved.

In addition, if achieving the outcome measures requires actions by both the provider and the patient, and if the provider only has limited influence over the patient's actions, then it may be appropriate to focus attention on the actions that are totally within the control of the provider versus actions the patient must take. For example, if a patient needs to take medications in order to achieve an outcome, the patient may be unwilling to do so, even if the provider removes the barriers.

On the other hand, use of process measures can be problematic if one of the key goals of the APM is to enable care to be delivered in different ways in an effort to achieve better outcomes. In most cases, the "evidence" that exists regarding the connection between processes and outcomes does not and cannot guarantee that carrying out a specific process (or achieving an intermediate outcome) will achieve a better outcome than all other alternative approaches.¹⁷⁴ Forcing a provider to use only the current "best practices" could preclude the development of "better best practices."

For the purposes of ensuring the accuracy and reliability of information used in the APM, such as the accuracy of diagnosis and the accuracy of outcome measures, process measures are likely to be essential. For example, for some diagnoses, the best way to assure the accuracy of the diagnosis will be to require that specific types of tests be performed and that those tests produce specific results. For an outcome measure based on patient-reported data, it will be important to ensure that the processes used to collect the data provide a valid, reliable measure.

iv. Striking the Right Balance Between Processes and Outcomes

The choice between process measures and outcome measures should be based on the goals of the APM and the care delivery model it is designed to support. Different choices will be appropriate in different situations:

- 1. Process measures will likely be most appropriate when the goal of the APM is to achieve more reliable or efficient delivery of current evidence-based processes than is possible under the current payment system.** For example, if the APM is paying primary care providers for proactive outreach to women in an effort to increase the proportion of women at risk of breast cancer who receive mammograms, then a process measure – whether an at-risk woman receives a mammogram – would be an appropriate quality metric. Although the true goal is to reduce the rate of death from breast cancer, the primary care practice cannot affect the incidence of breast cancer nor the success of treatment if a cancer is identified, so it is more appropriate to hold the practice accountable for the process it can control rather than the ultimate outcome for the patients.
- 2. A combination of process and outcome measures will be desirable when the goal of the APM is to deliver care in ways that are not supported by the current payment system.** In these cases, the process

measures would be used to ensure that changes are made in the aspects of care where changes are desirable and that changes are not made where evidence-based approaches should continue without changes. The outcome measures would enable a determination of whether the changes are having positive or negative impacts on the patients.

- 3. Outcome measures will be preferable when APM participants can control most of the factors that are likely to affect outcomes, and when care delivery needs to be highly customized to unique patient needs.** In these situations, there is a significant risk that process measures will limit the ability to provide the best care for each individual patient.

v. Assuring Diagnostic Accuracy

Assuring the accuracy of the diagnoses of a patient's conditions will be much more important under most Alternative Payment Models than it has been under the fee-for-service system. The ability to use condition-based payments instead of procedure-based payments, to stratify payments based on the patient's condition, and to risk-adjust utilization and quality measures depends on having accurate information about the nature of the patient's health problems and risk factors.

However, one of the reasons that diagnostic errors exist is because of the difficulty of determining an accurate diagnosis in many situations, and these same difficulties will arise in trying to verify the accuracy of the diagnosis.

- In some cases, the most definitive method of determining a diagnosis is to use a test that is very expensive or involves risks to the patient. For example, the "gold standard" test for determining the extent of coronary artery ischemia is to perform a cardiac catheterization and angiogram on the patient, but that procedure is expensive and it involves small but serious risks of injury and death for the patient. As a result, many patients appropriately receive a diagnosis of heart disease that is based on less definitive tests. It would be inappropriate for an APM to require use of a dangerous or expensive test in order to verify a diagnosis solely for the purposes of determining the patient's eligibility to participate in the APM.
- In other cases, there is no definitive way to determine a diagnosis; some diagnoses are established by ruling out all other alternative diagnoses, or by successfully resolving symptoms using the treatment focused on that disease. In these cases, it would be impossible to require verification of the diagnosis before treatment begins and it would be inappropriate to expect providers to deliver treatment without adequate payment to support it.

Consequently, the right way to measure diagnostic accuracy will likely depend on the specific diagnosis, and it may involve a combination of process measures (e.g., to verify that the appropriate tests were performed), intermediate outcome measures (e.g., to verify that the results of tests were consistent with the diagnosis), and/or outcome measures (e.g., assessing whether the symptoms are being resolved through use of the treatment designed for that diagnosis).

For diagnoses involving a high degree of professional judgment by a physician or other healthcare provider, an option is to require a “second opinion” by a different healthcare provider. However, this will involve additional expense – the provider of the second opinion will need to be compensated for their time and expertise, and there may be costs of additional or repeat testing – and it could involve delays in treatment for the patient. The costs would be lower if a second opinion is only sought for a sample of cases, but this would only help assure that the probability of accuracy was high, not that the diagnosis was accurate for every individual patient.

c. Obtaining Data to Assess the Quality of Care

As a practical matter, no matter which quality measures would be most desirable in theory, it will only be possible to use measures for which the necessary data can be obtained in an accurate, reliable, affordable, and timely way. This requires creating a precise definition of what data are needed to assess a particular aspect of quality, and then identifying an existing source of those data or developing a new method for gathering the data.

i. Using Data from Existing Quality Measures

In some cases, it will turn out that a quality measure already exists that does exactly what is needed for the APM. If so, then the data needed to assess quality in the APM may already be collected. However, even if current quality measures are not defined in a way that matches the needs of the APM, it may be possible to repurpose the data that are already being collected for one or more of these measures in order to define a different measure that does meet the needs of the APM.

For example, if an APM is intended to reduce spending on diabetes care without harming quality, one way to assess the quality of care would be to determine if individual patients’ HbA1c levels have changed under the APM. There is no current quality measure defined as “worsening of an individual patient’s HbA1c level,” and one cannot tell whether individual patients have gotten worse by examining changes in overall averages. However, a measure of the changes for individuals could be constructed because providers are already collecting data on patients’ HbA1c levels, and they would simply need to calculate the difference in levels for individual patients between two points in time.

ii. The Cost of Collecting Additional Data on Quality

If data that match the definition of quality needed for the APM are not collected currently, new or modified data will be needed. However, there is growing recognition

that there are significant costs associated with collecting data on the quality of care. Depending on the type of data needed, costs may be incurred by the provider of a service, by the patient, by the payer, and/or by other entities in order to collect and report the data needed to calculate the measure. If these costs are not paid for, either through the APM or other mechanisms, the data may not be sufficiently complete or accurate for use in the APM.

Consequently, in designing an APM, an explicit choice may be needed between (a) increasing the size of payments under the APM sufficiently to cover the costs of collecting the most appropriate data on quality (which could affect the business case for the APM) and (b) redesigning the APM so that a less extensive or expensive approach to quality measurement can be used. For example, as noted above, if the APM holds a provider accountable for the payer’s total spending on the patient, there is the potential for the provider to reduce spending on a wide range of services, and a wide range of data on quality would be needed to protect patients against inappropriate reductions in all of those services. In contrast, if the APM is designed to focus accountability on spending for specific services, there would be fewer ways in which quality could be harmed and a more limited set of data would be needed to assess that.

Although using data that providers are already collecting may be less costly than defining and collecting new types of data, it may be preferable to incur the costs of collecting data for a new, correctly-specified measure than to jeopardize patient safety or to penalize providers inappropriately by using existing but inappropriate data to measure quality. Because the APM will be paying differently for care delivery, it could also be designed to pay adequately to support the collection and verification of important quality data that are not currently being collected.¹⁷⁵ Alternatively, it may be possible to eliminate existing requirements for the provider to collect other, less relevant types of quality data, thereby freeing up time and resources to collect the new data.

It is important to recognize that if data on a specific aspect of quality are being collected for the first time as part of an APM, it may not be possible to use those data for accountability purposes for a period of time until it is clear that the data are valid and reliable, and/or until sufficient data are available to establish baseline levels of performance. Consequently, the initial period of implementation of the APM may need to use a narrow or temporary set of measures for accountability purposes until the data for the full, desired set are available. Many CMS APMs have used “pay for reporting” rather than “pay for performance,” during the initial year of implementation, i.e., making payment contingent on the provider reporting data on quality but not contingent on the level of performance achieved.

2. Setting the Performance Targets for Quality

Once the key aspects of quality that can be affected by the APM and methods of obtaining data to assess them have been identified, a decision must be made regarding the level of quality the APM participants will be expected to achieve in each of those areas. In order for a payment model to qualify as an APM, the quality of care must not be worse than it would have been otherwise, and if the APM does not reduce spending, it must improve quality compared to what it would have been otherwise.

a. Patient-Level Targets vs. Population-Level Targets

Existing APMs typically evaluate changes in the quality of care based on changes in population-level quality measures that calculate the percentage of patients for whom a process was performed or a particular outcome level was achieved. In order to determine whether quality is better or worse than it would have been otherwise, the percentage is compared to a previous period or to patients who are not participating in the APM.

There are several problems with this approach:

- **The definition of “quality” used in standard population-level quality measures may not assess the specific types of impacts likely to be associated with the APM.** All population-level quality measures are based on some type of patient-level target for quality, e.g., a care process that is to be performed, or an outcome level that is to be achieved. However, that target may not be appropriate for the specific patients who are in the APM, or it may be insensitive to the changes in quality that are likely to result from the APM. For example, a quality measure that is used in many APMs is the percentage of diabetic patients whose glycated hemoglobin (HbA1c) level is above 9%. If an APM is designed to improve care for diabetic patients who have multiple comorbidities and lifestyle challenges and for whom standard medications have been unsuccessful, a reduction in their average HbA1c levels from 12% to 10% would likely be viewed as a successful improvement in diabetes control, yet the standard quality measure would show no change.
- **The quality of care may have worsened for individual patients even though the measure showed no change or even improvement.** Most population-level measures are based on whether a process was performed or an outcome was achieved during a particular period of time for the patients who were receiving care during that period of time, without regard to the quality of care the same patients received during the previous time period. If there is no change in the measure from one period to the next, that could be because every patient received the same quality of care as they had in the previous period, but it could also be because some patients received worse care while other patients received better care. For example, if one-tenth of diabetic patients experienced a reduction in HbA1c levels from 9.5% to 8.5%, one-tenth of diabetic patients experienced an increase from 8.5% to 10.0%, and the remaining patients con-

tinued to have HbA1c levels below 9%, the percentage of patients with good control would appear to have remained unchanged, even though one-tenth of the patients did worse under the APM. In addition, depending on how many patients are participating in the APM, small reductions in quality or even large reductions in quality for small groups of patients might not be visible in the measure.

- **The patients used for comparison purposes may differ from the patients in the APM in ways that mask changes or differences in quality.** Most population-level measures are not risk-adjusted based on characteristics of patients that can affect the quality of care. If the patients participating in the APM differ in significant ways from the patients used as a comparison group, any difference or lack of difference in a quality measure may reflect differences in the patients instead of or in addition to any impacts of the care related to the APM.

From a patient’s perspective, what matters is whether the APM is maintaining or improving the quality of care *that individual patient* receives, not what the APM does for *other* patients. Consequently, the starting point in setting quality targets for an APM is to define appropriate *Patient-Level Targets*, i.e., the threshold(s) that will be used for determining if an individual patient is benefiting or being harmed by participating in the APM.

Then, if appropriate, a *Population-Level Target* can also be defined to assess the extent to which the Patient-Level Targets are being achieved for a population of patients, e.g., all of the patients who are participating in the APM. Although it is problematic if any individual patients are being harmed by participation in the model, it is not necessary that every patient receive improved care in order for the APM to be deemed successful, just as an APM can be successful financially if savings are achieved for some but not all patients.

In some cases, it may be possible to use the specifications of existing quality measures in order to define Patient-Level Targets and to evaluate population-level changes in quality for an APM. However, in many cases, existing quality measures will be inadequate for the purposes of an APM and so they will either need to be modified in some way or completely new quality measures will be needed.

b. Defining Patient-Level Targets for Quality

There are several approaches that can be used to define Patient-Level Targets for quality, including:

- maintaining the prior level of quality for the individual patient;
- achieving evidence-based standards or guidelines;
- achieving statistically significant improvement;
- achieving clinically important improvement; and
- achieving patient-specific goals.

i. Prior Quality Levels for Participating Patients

In some cases, many or all of the patients who are receiving services supported by the APM will have previously been receiving services for the same condition or issue. This would include patients who are being treated for a chronic or long-term illness and patients who are receiving preventive or wellness care. In these situations, the goal of the APM should be to deliver care to these patients that is no worse than, and ideally better than the care the patients received previously. Consequently, the Patient-Level Target for each patient should either be (1) the level of quality or outcomes achieved for that patient immediately prior to the period of time during which performance in the APM is being measured, or (2) some defined increment of improvement beyond that previous level.

Setting Patient-Level Targets in this way when the APM is first implemented requires baseline data on the quality of care for the specific patients who participate in the APM:

- If the patients are receiving care from the same physicians, hospitals, or other providers under the APM as they were before, and information on the relevant aspect of quality has been collected by the providers in the past, then it should be feasible to set Patient-Level Targets for each individual patient using those data.
- If the patients are insured by the same payer and the relevant aspect of quality has been measured and reported to the payer at the patient level in the past, then it should be feasible to use the payer's data on prior levels of quality to set Patient-Level Targets.
- If no prior records are available, or if the patient has not previously been treated for the condition, it may be possible to establish a baseline level by assessing each patient before they begin treatment under the APM.

If baseline data on the quality measures are not available, then a different method of setting Patient-Level Targets will be needed, at least initially. After the APM is in place, data on the quality measures will presumably be collected on an ongoing basis, so Patient-Level Targets in the second and subsequent performance periods could be based (in whole or in part) on maintaining or improving quality compared to the initial period.

If the patients have a condition that normally worsens over time even with the best quality care, or if new types of treatment are introduced that improve care for all patients with the condition, then adjustments to baseline quality/outcome levels may be needed in order to set Patient-Level Targets that more realistically reflect what would have happened in the absence of the APM. This is analogous to creating counterfactual benchmarks for spending accountability.

Setting Patient-Level Targets based on the prior level of quality patients received can be used to assure that the quality of care does not decrease under the APM, but that may or may not mean that the quality of care is *good*. That could be addressed by using evidence-based standards (as discussed in the next section) to establish

Patient-Level Targets instead of or in addition to the baseline quality of care the patients have been receiving.

However, it may be unreasonable for an APM that is focused on changing one aspect of care to do more than maintain the current level of quality in other areas. Requiring providers in the APM to make improvements on every aspect of quality could discourage their participation, thereby causing patients to lose the opportunity for any improvement in quality. For example, if an APM is designed to enable pulmonologists to improve care for chronic obstructive pulmonary disease (COPD), some of the patients with COPD may also have heart failure and/or diabetes. One would not want a pulmonologist to improve treatment of these patients' COPD in a way that makes their diabetes or heart failure worse (as could happen depending on the types of drugs used to treat their COPD), so it might be appropriate to include quality measures related to diabetes and heart failure and use them to ensure that quality has not worsened. However, it would not be reasonable for the APM that is focused on COPD to also require that the pulmonologist *improve* the care of the patients' diabetes and heart failure. If patients with COPD, diabetes, and heart failure are receiving poor quality care for all of the conditions, then it may make sense to create an APM specifically designed to enable a team of physicians to improve the quality of care for all three conditions in a coordinated way.

ii. Evidence-Based Standards or Guidelines

If research has shown that patients who receive a particular process or reach a specific level on an intermediate outcome measure consistently do better in terms of true outcomes, then delivery of that process or achievement of the intermediate outcome could be used as a Patient-Level Target. This approach could be used for acute conditions and newly diagnosed chronic conditions, as well as for chronic condition care and preventive care that a patient had been receiving prior to implementation of the APM.

If the evidence indicates that different processes or intermediate outcome levels are appropriate in order for different patients to achieve similar ultimate outcomes, then the Patient-Level Targets would need to be different for those different patients. If evidence is only available for a subset of patients, then an evidence-based threshold could be used for those patients, and thresholds for the other patients could be established using one of the other methods.¹⁷⁶

It is important to note that even if Patient-Level Targets are set using evidence-based standards, the quality of care for some patients could still be worse under the APM than it was before. If the standards define a minimum level of quality that must be achieved, and if care quality exceeded that minimum prior to the APM, it is possible that the APM could achieve savings by reducing quality to the minimum. Conversely, the quality of care for some or all patients could be better than it was in the past or better than what it would otherwise have been even though it still falls short of what evidence indicates would be desirable, particularly if the APM does not or cannot remove all of the barriers to achieving evidence-based care. Consequently, it may be ap-

appropriate to set Targets based on whichever is better – achieving the evidence-based standard or maintaining the quality level achieved for the patient in the past.

iii. Statistically Significant Improvement

In many cases, the evidence is not sufficiently strong or clear to enable anyone to say with confidence that a particular process or outcome is desirable or undesirable for particular patients. Most guidelines and appropriateness criteria include a category of patients for whom the desirability of delivering a particular process or achieving a particular outcome is “uncertain,” and this group – the gray area – may include the majority of patients.¹⁷⁶ Conversely, even if clear evidence exists that a particular threshold results in the best outcomes, there may be barriers to achieving that threshold for the patients participating in the APM that did not exist in the research that produced the evidence, and the APM may not have been designed to fully overcome those barriers.¹⁷⁷

In these cases, if the goal of the APM is to achieve an improvement in care quality or patient outcomes, there will need to be a different way of defining how large the improvement must be. Since most measures of quality or outcomes will have some degree of random variation in their values even if no true change has occurred, it will be desirable to ensure that any measured change is not merely a result of random variation. This can be done by assessing the “normal” variation in the measure (using either multiple measurements for the individual patient if such data are available or measurements for multiple similar patients) and requiring that an improvement be large enough to represent a “true” change at some pre-specified level of statistical confidence.

iv. Minimal Clinically Important Difference

The fact that a measure of quality or outcomes improves, even if the improvement is greater than what would likely be expected from random variation, does not mean that the improvement is large enough to have any real value to the patient. To address this, the Target could be defined as achieving at least a “Minimal Clinically Important Difference.”¹⁷⁸ For example, if a patient with heart failure can only walk 10 feet on the standard 6-minute walk test before experiencing chest pain or shortness of breath, a 20% improvement (an additional 2 feet) would be unlikely to enable the patient to perform any important activities they cannot currently perform. A much larger change (e.g., the ability to walk 100 feet) would be needed to enable the patient to recognize that improvement has occurred or to enable the patient to adequately perform activities of daily living or to engage in recreational activities.¹⁷⁹

v. Patient-Specific Goals

Even if patients achieve a level of improvement sufficient to be deemed a Minimal Clinically Important Difference, that “minimal” level may not be viewed as sufficient to justify participation in the APM by the payer or patient. In these cases, the Patient-Level Target could be a specific goal established by the payer or patient, or a goal established by the APM participant that the pa-

tient and/or payer find acceptable. Such Patient-Level Targets could be established in terms of either absolute levels or for improvements relative to a baseline. For example, if the APM is intended to reduce the cost of hip or knee surgery, one of the Patient-Level Targets could be for the patient to experience minimal pain after recovery from the surgery. If the APM is intended to redesign hip or knee surgery to enable patients to walk farther without pain than before the surgery, the Target could be a specific additional distance or percentage increase in distance relative to each patient’s baseline level needed for the patient to carry out specific tasks or participate in specific activities.

c. Defining Population-Level Quality Targets

i. When and Why Population-Level Targets Are Needed

An APM cannot be considered successful if individual patients are being harmed as a result of a change in care delivery that is required or encouraged by the APM. Consequently, if a Patient-Level Target has been defined in terms of the previous level of quality for the individual patient, or in terms of a minimum acceptable level of quality, then failure to achieve the Target for any individual patient represents a failure by the provider that is caring for that patient and it potentially represents a failure of the APM as whole. In these cases, the Population-Level Target would need to require that 100% of the patients achieve the Patient-Level Target.

However, for patients with acute conditions or newly-diagnosed chronic conditions, there will be no prior level of quality for that specific patient, so Patient-Level Targets will need to be defined based on evidence-based standards or based on goals. Although it would be ideal if every patient achieved these Targets, failure to achieve the Targets for some patients could still represent an improvement in care if (a) a higher proportion of patients had failed to meet similar Targets in the past or (b) patients had previously fallen even farther short of the Target than under the APM. A Population-Level Target is needed to assess whether or not this occurred.

If the goal of the APM is not to reduce spending, but to improve quality, or if the goal is to improve quality as well as reducing spending, it would not be necessary for every patient to receive better quality care. As long as no patient received *worse* care, even if only a subset of patients received *improved* care, the overall quality of care for the group of patients will have improved. This is particularly true for many types of outcome measures. For example, although knowledge about how to treat cancer is expanding rapidly, in many cases, only a fraction of patients achieves good outcomes even when treatments are based on the best evidence available, so if an APM enables a higher percentage of patients to achieve good outcomes, that would represent improvement. Similarly, although significant progress has been made in reducing many types of hospital-acquired infections, it is rare to find any hospital that has completely eliminated them on a sustained basis, and so a reduction in the rate of infections would represent a success.

Consequently, if the APM is expected to improve quality, two sets of Targets should be defined:

- a Patient-Level Target that defines the minimum level of quality that must be achieved for each patient, with an associated Population-Level Target of 100% success in achieving that Patient-Level Target; and
- a second Patient-Level Target that defines the higher-than-minimum level of quality that is desired for each patient, and an associated Population-Level Target defining the proportion of patients who need to achieve that higher level in order for the APM to be viewed as successful.

ii. Alternative Forms of Population-Level Targets

Some aspects of quality can only be good or bad, with nothing in between. For example, an essential service is either delivered to the patient or not, and a patient either dies or does not. “Never events,” such as wrong-sided surgeries, either happen or they don’t, and a never event is viewed as undesirable in all situations. For these aspects of quality, the Population-Level Target can be defined in terms of the *percentage of patients who achieved the Patient-Level Target*. When poor quality occurs only rarely on a particular measure, the measure is typically expressed as a rate rather than a percentage, e.g., if 10 out of a group of 5,000 patients experienced poor quality care, this would be expressed as a rate of 2 per thousand, rather than 0.2%.¹⁸⁰

In other cases, quality is inherently continuous rather than binary. For example, pain is not only present or absent, but it varies in severity. A patient’s ability to walk without pain after completing rehabilitation could vary from a short distance to a long distance, and their blood pressure following treatment could vary from “too low” to “slightly high” to “very high.” For aspects of quality that are continuous, Population-Level Targets based on percentages alone fail to accurately measure the true nature of the quality of care that is being delivered.

Many current quality measures convert a continuous quality measure into a binary measure by assessing whether a patient is above or below a specific quality threshold. However, that makes no distinction between patients who are close to the threshold and those who are far from it. For example, a commonly used measure of the quality of care for diabetic patients is based on the percentage of patients with an HbA1c level lower than 7%, but the measure makes no distinction between patients with an HbA1c level of 7.1% and a patient with a level of 7.9%, even though the latter patient might be at much higher risk of diabetic complications.

If it matters not just *whether* the quality of care fell short of the Patient-Level Target, but by *how much*, then one or more additional Population-Level Targets for the *distribution* of quality across the population may be needed. These Targets could be defined using measures of the characteristics of a distribution, such as the average deviation from the Patient-Level Target. For example, if the Patient-Level Target for HbA1c levels in a group of patients is 7%, then in addition to setting a Population-Level Target for the percentage of patients who had an HbA1c level below 7%, one might also establish a second Population-Level Target in terms of the average dif-

ference between the actual HbA1c level and 7%. A provider participating in an APM whose patients routinely had HbA1c levels much higher than 7% could then be penalized more than a participant whose patients had HbA1c levels only a small amount above 7%. Current quality measures for diabetes attempt to assess the distribution by calculating the percentage of patients with an HbA1c higher than 9% as well as the percentage lower than 7%, but this makes no distinction between a group of patients with an HbA1c level of 7.1% and a group of patients with a level of 8.9%.

Before defining Population-Level Targets in either percentage or distributional terms, it is important to ensure that the Patient-Level Target is appropriate for each of the patients. If the denominator of the percentage includes patients for whom the Patient-Level Target is not applicable, then a percentage less than 100% does not necessarily indicate that some patients received poor quality care. A problem with many current quality measures is that patients for whom the Patient-Level Target was inapplicable are counted the same way as patients who should have achieved an applicable Patient-Level Target but didn’t.

If it is desirable and appropriate that processes or outcomes of care differ for some patients, then the Patient-Level Targets should differ for those patients, and the Population-Level Target could assess the extent to which the appropriate Patient-Level Target was met, using either percentage or distributional measures. For example, current guidelines for management of diabetes call for HbA1c goals to be customized, so to support that, each patient should have their own Patient-Level Target, and the Population-Level Target would be for every patient to achieve their own Patient-Level Target.

Alternatively, if there are two or more well-defined subcategories of patients, and if all patients in a subcategory have the same Patient-Level Target, then the Population-Level Targets could be *stratified*, i.e., a separate Population-Level Target would be established for each subcategory of patients. If the payment amounts or the utilization/spending measures have been stratified into separate categories to reflect important differences in patient needs, then it will likely be desirable to also stratify the Population-Level Targets for quality into the same categories.

iii. Alternative Ways of Setting Population-Level Targets

Similar to the methods of setting Population-Level Targets for utilization and spending that were discussed in Section VI.B, there are three fundamentally different ways of setting the Targets for quality:

a. Status Quo-Based Targets

Since an APM must not make the quality of care worse than it would otherwise be, and it may be expected to improve the quality of care, one approach is to define a Population-Level Target based on an estimate of what the quality of care would have been in the absence of the APM. In the context of spending, this is typically referred to as a benchmark, but in the context of quality, the term “benchmark” ordinarily connotes the best

level of quality that has been achieved in the past or by others. Consequently, the term “status quo” will be used here to represent the level of quality that would have occurred if the APM did not exist.

As with Benchmark-Based Targets for utilization and spending, Status Quo-Based Targets for quality have two components:

- an estimate of the Status Quo level of quality, i.e., the level of quality that is currently being achieved; and
- a Target Change, i.e., the amount, if any, by which quality must improve compared to the Status Quo in order for the APM to be viewed as successful.

b. Evidence-Based Targets

If there is evidence regarding the quality of care or outcomes that can consistently be achieved for the types of patients participating in the APM when they receive the services the APM is designed to support, then those levels of quality or outcomes could be used to define the Quality Targets for the APM.

c. Competitive Targets

If there are multiple providers participating in the APM, the Target could be set through a competitive process, i.e., each provider would define the Patient-Level and Population-Level Targets for each aspect of quality that it was willing to be held accountable for achieving, and then payers or patients could choose a provider based in part on a comparison of their Target Levels to those of other providers, as well as a comparison of their prices for services and the Targets they had established for utilization and spending. Regardless of any explicit financial penalty or reward in the APM based on quality, if patients valued the higher level of quality delivered by a provider and could choose to receive services from that provider, the provider would be rewarded with more patients and more revenue. If patients begin switching to providers that committed to achieve higher Targets, other providers would have an incentive to also commit to the higher Targets in order to retain patients and attract new patients.

This process would encourage higher-value healthcare delivery using the same kinds of market forces that encourage development of higher value products and services in other industries. Minimum Quality Targets could be required from all APM participants, but individual APM participants could voluntarily commit to higher Quality Targets if they believed they knew how to redesign care in a way that would achieve higher levels of quality. An individual patient could make choices between providers participating in the APM based on a clear understanding of the level of performance that each provider would achieve for *that* patient, rather than based on the provider’s historical average performance for *other* patients.

iv. Methods of Defining Status Quo-Based Targets

If Status-Quo Based Targets are used, there are at least two different ways of defining the “status quo:”

a. Maintaining or Improving on Prior Performance

If the patients participating in the APM have received care for the same conditions in the past, then the quality of care they received previously could be used to define the Status Quo level that must be maintained or improved in the future. For most types of acute conditions and for newly diagnosed or treated chronic conditions, however, the patients who participate in the APM will not have previously been treated for the condition, so there will be no baseline for those specific patients. In these cases, the Status Quo level could be the level of quality the provider achieved for similar patients in the past.

The ability to use this approach depends on whether providers collected the necessary data on the quality measures during the previous period. In addition, if the patients participating in the APM are different from other patients in ways that would affect the way care was delivered or the outcomes achieved, then a Target set using this approach could be inappropriately high or low unless there is an adequate way of adjusting for those differences.

b. Maintaining or Improving on Performance Achieved by a Comparison Group

An alternative approach is to define the Target as the level of quality achieved for a similar group of patients who are treated by providers that are not participating in the APM. This would help ensure that if the quality of care or outcome is improving for patients outside of the APM, a smaller improvement inside the APM would be treated as a failure rather than a success. Similarly, if the quality of care or outcomes are worsening for other patients, the APM participants would not be penalized for experiencing a similar problem.

In the context of quality measures, this is what is typically referred to as a “benchmark.” (As noted previously, this use of the term “benchmark” is different from the usage of the term that has become common for evaluating spending.) For example, the Achievable Benchmarks for Care (ABC™) methodology determines the minimum performance level that has been achieved for at least 10% of patients by the highest performing subset among a group of providers.¹⁸¹

There are two important challenges in using this approach:

- **Identifying a comparison group that is similar to the APM patients.** It may be difficult to find a similar group of patients, or to verify they are similar, if the eligibility criteria for the APM are based on information that is not routinely collected by all providers or if there is a reason to expect that patients with specific characteristics will self-select into or out of the APM;
- **Obtaining the quality/outcomes data on the comparison group.** The data needed to measure quality or

outcomes in the comparison group may not be collected by the providers who are caring for those patients, particularly if the quality/outcome measure has been specifically developed for the APM, or if the costs of collecting the data are high.

Consequently, this approach may only be feasible for a subset of APMs or if a separate initiative is created to obtain the necessary data on patients who are not participating in the APM.

v. Defining Target Changes

If there is an expectation that quality will *improve* on the measure, then in addition to determining the Status Quo, a non-zero Target Change will need to be defined. Four approaches that can be used for this are:

- **Goal-Based Change.** The goal might be based on the level of improvement that would be viewed as sufficient by either payers or providers to justify implementing the APM. In contrast to a goal-based patient-level target, the goal for a population-level measure could be to achieve improvement for a minimum percentage of patients.
- **Statistically Significant Change.** If it is not clear how much improvement is possible, or if there is significant random variation in quality levels, an alternative is to define the minimum level of change needed to provide assurance that real change has occurred rather than what appears to be a change that is merely random variation. Requiring a statistically significant change at the population level is a weaker standard than requiring a statistically significant change for individual patients, because large improvements for some patients could offset smaller improvements or even worsening of quality for other patients, such that on average, there is a statistically significant improvement for the overall population.
- **Clinically Important Difference.** As with patient-level targets, the fact that a statistically significant change has occurred does not mean that the change is large enough to be meaningful to patients. To address this, the level of change required would be the minimum needed to be perceived by patients as an improvement in one or more outcomes. In contrast to the Patient-Level Target, the Population-Level Target could require that enough patients achieve clinically meaningful improvements to offset any reductions in quality or outcomes for others.
- **Comparison Group Change.** The Target Change could be defined to be equal to or better than the change in quality for a comparison group not participating in the APM.

d. Issues in Defining Targets

Many of the issues described in section VI.B with respect to Targets for utilization/spending measures also apply to the Targets for quality measures.

i. Prospective vs. Retrospective Targets

One of the problematic aspects of some APMs and pay-for-performance programs is that the quality Targets that providers are expected to achieve are not defined until after patients have already received the care, i.e. they use *Retrospective Targets*. For example, in the Medicare Hospital Readmissions Reduction Program, hospitals receive penalties based on their performance on readmission rates, but the performance standard is determined by how *other* hospitals perform during the *same period of time*. This means that even if a hospital performs well compared to how other hospitals performed in the past, or even if the hospital improved compared to its past performance, the hospital could still be penalized if it turns out that most other hospitals improved even more.

In contrast, *Prospective Targets* for quality enable the APM participant to clearly understand the level of performance that is required *before* care delivery begins. This enables the participant to monitor performance and make adjustments in care delivery if performance is falling short of the target. The methods described earlier for setting Patient-Level Quality Targets all result in Prospective Targets: If the Quality Targets are based on goals or evidence-based standards, those goals or standards are known before care begins, and the Quality Targets would only be based on the quality of care the patients received in the past if that information is available prior to the delivery of care.

A Population-Level Quality Target can also be prospective if it is established competitively or is based on the provider's past performance. However, if the Population-Level Target is based on the performance of a comparison group during the same performance period, it would be impossible to define the Target until after the performance period has ended.

ii. Participant-Specific Targets vs. Common Targets

Similar to the discussion in Section VI.B with respect to Utilization/Spending Targets, an APM can either use the same Quality Target for every APM participant or create a customized Target for each participant. For example, some APMs are designed to reward "improvement" instead of or in addition to "achievement;" this means that if one APM participant had lower quality performance in the past than another, the Target for the first participant would also be lower. Both participants might or might not be expected to achieve the same level of improvement from the past, but the resulting level of quality would be different.

Participant-Specific Targets can encourage participation in the APM by a larger number of providers if it makes the Target easier to achieve for lower-performing providers. However, they can be problematic from the perspective of patients, since they can result in the same

amount of payment for what appear to be two very different levels of quality.

If some providers have lower levels of performance on a quality measure than others because of greater challenges in achieving high quality for the specific types of patients they treat, then rather than setting different Targets, it would be preferable to stratify the Targets by patient characteristics and use the same Target in each category for all APM participants.

iii. Revising Targets and Changing the Target Methodology Over Time

Once a Target is set, a decision must be made as to how often the Target will be revised and in what way it will be revised. If new evidence emerges about what types of care processes are necessary to achieve good outcomes, or if research shows that certain treatments have undesirable effects on specific patients, then it will be important to revise the Quality Targets appropriately so that the APM is not rewarding bad care or penalizing providers for delivering what is best for patients.

Over time, it may also be necessary to change the methodology for setting Targets. For example, if a majority of providers are participating in the APM, it will be difficult to find comparable providers not participating in the APM that can be used for setting Targets based on a comparison group.

**TABLE 11
ALTERNATIVE WAYS OF SETTING TARGETS FOR QUALITY**

Alternative	Strengths	Weaknesses	Examples
I. Patient-Level Target	Ensures that quality is maintained or improved for each patient	Assumes that care can be designed to address the needs of each patient	
A. Prior Level of Quality for the Patient	Ensures quality is maintained or improved from what patient has experienced previously	Prior level of quality may have been unacceptably low or anomalously high	Patient can walk without pain or shortness of breath the same distance as previous year
B. Evidence-Based Standard	Ensures the quality goal is feasible to achieve	Patient may have already been achieving better outcomes; Evidence may not exist for how to feasibly and reliably achieve good quality for similar patients	Patient can walk without discomfort for a distance similar to what patients could walk in a controlled trial following the type of treatment supported by the APM
C. Statistically Significant Difference	Ensures that a change in quality is not due solely to random variation	Penalizes small providers because significance is primarily driven by number of patients	The distance the patient can walk increased by more than normal random variation
D. Clinically Important Difference	Ensures the change in quality has a meaningful impact on patients	The change may not be large enough to justify implementing the APM	The patient feels they are better able to perform activities without pain or shortness of breath
E. Patient-Specific Goal	Ensures that change in quality is important to the patient or payer	The goal may not be feasible to achieve	The patient has the ability to adequately perform all activities of daily living without pain or shortness of breath
II. Population-Level Target	Allows quality to be improved for some but not all patients	Could result in lower quality for a subset of patients if quality improves sufficiently for others	
A. Evidence-Based Target	Ensures the quality goal is feasible to achieve	Patients may have already been achieving better outcomes; Evidence may not exist for how to feasibly and reliably achieve good quality for similar patients	Rate of hospital-acquired infections (HAIs) repeatedly achieved in a controlled trial using a standard infection prevention protocol
B. Competitive Target	Encourages innovation in achieving higher quality	Requires multiple providers to compete based on quality/outcomes; May or may not represent improvement over current levels of quality	Lowest rate of HAIs achieved by other APM participants
C. Status Quo-Based Target	Helps ensure quality is not below current levels	May be lower quality than what is feasible to achieve	
1. Status Quo Definition			
a. Prior Performance for Same Patients	Allows quality Target to be based on unique characteristics of the patients	Cannot be used for new acute conditions; Patients with chronic conditions may have received poor quality care previously	Rate of HAIs during previous hospitalizations for the patients (e.g., patients with a chronic disease who are frequently hospitalized)
b. Prior Performance for Patients of Provider	Allows a provider-specific Target to be used for providers treating acute conditions	Will result in lower-quality Targets for providers who have had lower quality in the past	Rate of HAIs for similar patients treated by the provider for the same condition
c. Comparison Group	Ensures quality will be no worse than expected in the absence of the APM	Patients in the APM may have different characteristics than those not participating	Rate of HAIs among similar patients not participating in the APM

(CONTINUED)

**TABLE 11 (CONTINUED)
ALTERNATIVE WAYS OF SETTING TARGETS FOR QUALITY**

Alternative	Strengths	Weaknesses	Examples
II. Population-Level Target			
C. Status Quo-Based Target			
2. Target Change			
a. Goal-Based Change	Ensures that the APM achieves results that justify the effort in implementing it	May achieve lower quality than is possible; Goal may not be feasible to achieve	Elimination of HAIs
b. Statistically Significant Change	Ensures the improvement was not due solely to random variation	Change could be very small if number of patients is large enough; Penalizes providers with small numbers of patients; Cannot be used for low frequency events	Statistically significant reduction in HAIs
c. Clinically Important Difference	Ensures the change in quality has a meaningful impact on patients	Impact may not be large enough to justify implementation of APM	Reduction in number of HAIs that result in death or permanent disability
d. Comparison Group Change	Ensures quality will be equal to or better than what is achieved for similar patients not participating in the APM	Could permit a reduction in quality for the patients in the APM if quality is decreasing elsewhere; Does not define how much quality should improve if the APM is intended to achieve improvements; Patients in the APM may have different characteristics than those not participating	Change in HAIs equal to or better than the change achieved for similar patients not participating in the APM

3. Assessing Performance on Quality

As with accountability for utilization/spending, assessing an APM participant's performance on quality is generally not as simple as comparing the participant's actual level of quality in a particular period of time with the Quality Target that is defined for that time period. The assessment methodology needs to determine the extent to which any difference between the quality measure and the Target for that measure were due to the APM participant's performance or due to other factors.

There are five basic reasons why an APM participant's actual performance on quality can exceed or fall short of the Target:

1. **Performance Success or Failure**, i.e., whether the APM participant took the necessary actions to maintain or improve quality as had been expected under the APM.
2. **Effects of Errors in Calculation or Measurement**. The measure of quality or the Target could be computed inaccurately due to errors in the data collected or the calculations made.
3. **Effects of Inadequate or Inaccurate Adjustment for Uncontrollable Factors**, i.e., the quality measure or the Target did not correctly or completely adjust for known factors that affect quality but are not controllable by the APM participant.
4. **Effects of Rare and Unpredictable Events**. Some circumstances that affect quality or outcomes will occur infrequently and unpredictably, e.g., a rare variant of a disease or an unusual confluence of circumstances that cause complications of treatment.
5. **Effects of Random (Unexplained) Variation**. Differences among specific patients may cause variations in quality for reasons that cannot be predicted and cannot be controlled by the APM participant.

The extent to which the difference between an APM participant's quality of care or outcomes and the Target truly reflects the participant's success or failure depends on the relative size of the first factor versus the others.

Similar to the discussion in Section VI.B, random variation will be a more important factor when the APM participant is treating a smaller number of patients, when the patients are more diverse, and when performance is being measured over a shorter period of time. In addition, the most important quality measures typically involve rare events such as deaths, and it may be impossible to reliably calculate such measures or assess performance for small groups of patients and/or short periods of time.

Statistical methods can help in estimating the extent to which the APM participant's actions contributed to the results, but the uncertainty caused by random variation means that errors in these estimates are unavoidable. If the Target is based on ensuring there is no reduction in quality, then it would be a "Type I error" to conclude that there had been a true change in the quality of care delivered by a particular provider when the change in the quality measure was due to random variation or error. Conversely, it would be a "Type II error" to conclude that

there had been no change in quality simply because the change *might* have been due to random variation.

Similar to the discussion in Section VI.B with respect to assessing performance on utilization/spending, a decision has to be made about the balance between Type I and Type II errors in assessing performance on quality. A provider should not be penalized because it appears that quality worsened or failed to improve due solely to random variation, but it is inappropriate to allow patients to continue to be harmed by the care delivered under an APM simply because the standard for proving harm is too stringent. Bayesian methods can be used to create greater certainty in the evaluation of quality performance. For example, Bayesian methods are currently used by CMS for reporting differences in mortality rates across hospitals.

Both Type I and Type II errors are less likely if the amount of normal variation among patients and the magnitude of any measurement errors is low, since it will be easier to determine that a small change in the quality measure represents a true change in quality. Consequently, if there is a choice of measures for assessing a particular aspect of quality, errors will be smaller with the measure where the data can be most accurately collected and where there are fewer ways in which the measure can be affected by factors outside of the control of the provider. A process measure will often be more likely to have these characteristics than an outcome measure, so in this respect, process measures can be preferable to outcome measures for accountability purposes, assuming that evidence shows a close connection between delivery of the processes and achievement of the desired outcome.

4. Making Performance-Based Adjustments to Payments

Once the quality measures, targets, and methods of assessing performance are defined, the final step is defining the mechanism by which APM participants will be penalized or rewarded based on how actual performance compares to the targets. The options are similar, but not identical, to those described in Section VI.B for utilization/spending:

1. Penalties, or penalties and bonuses, based on whether Quality Targets are achieved, in addition to payments for services delivered to patients.
2. Paying for services to patients only when Quality Targets are achieved, i.e., “outcome-based payments;”
3. Warranted payments for services;
4. Terminating a provider’s participation in the APM if the provider fails to achieve the Quality Targets; and
5. Terminating the APM if providers fail to achieve the Targets.

Option 1: Penalties/Bonuses in Addition to Service-Based Payments

Under this option, the healthcare provider that is participating in the APM is paid for delivering the desirable services using whatever methodology is defined for Component #1, but the APM participant is required to pay a penalty if the Targets on one or more quality measures are not achieved.

This option requires a methodology for determining the magnitude of the penalty that will be imposed if the Target for the quality measure is not achieved. The larger the penalty, the greater the financial incentive for the APM participant to achieve the desired result, but the greater the potential financial problems the participant could face, particularly if the penalty represents a large proportion of the participant’s revenue.

a. Patient-Level Targets vs. Population-Level Targets

Penalties can be used with either Patient-Level Targets or Population-Level Targets:

- If a Patient-Level Target is not met for an individual patient, the provider would pay a penalty for that patient, e.g., by charging the patient less or refunding all or part of the payments that the patient or their payer had already made.
- If a Population-Level Target is not met, the provider would either pay a penalty to the payer for a group of patients or would pay each individual patient a share of the overall penalty.

An important weakness of using Population-Level Targets under Option 1 is if the provider participating in the APM achieves the Target level of quality, the provider would pay no penalty, but the quality for a subset of patients could be worse than it would have been otherwise. If those patients are paying for their own care, or if their health insurance requires cost-sharing for services (whether it be in the form of a deductible, co-insurance,

or a copayment), those patients would be paying for lower-quality care under the APM or paying for care that failed to achieve the outcomes that were expected. From the patient’s perspective, this is the opposite of how a value-based payment should work – the patient with a bad outcome should be paying less than the patient with the good outcome.

b. Making the Penalty Proportional to Performance

Instead of a fixed penalty for failing to achieve the Target, the penalty can be made proportional to the amount by which the quality measure fell short of the Target. For example, if the Target was for 100% of patients to be pain-free following the procedure, and if one provider achieved a 90% rate and another provider achieved an 80% rate, a higher penalty could be imposed on the second provider. This approach is particularly helpful when it is difficult to predict the exact impact of a service or where there is random variation in the desired result, since it avoids imposing a large penalty if the result falls just short of the Target level.

Example: In the CMS Comprehensive Primary Care Plus initiative, participating primary care practices must repay Medicare up to one-half of the Performance-Based Incentive Payment (PBIP) based on their performance on a set of quality measures.¹⁸²

c. Determining the Absolute Amount of the Penalty

When an APM participant fails to meet a Spending Target, the gap can be directly measured in dollars, and penalties can be proportional to those dollar amounts, as discussed in Section VI-B. However, when an APM participant fails to meet a Quality Target, it will not be as obvious how large the financial penalty should be. Although it will make sense to assign a larger financial penalty to a more serious quality problem (e.g., death as opposed to pain), a decision must be made about the absolute dollar penalties for some aspects of quality in order to scale the penalties for the others.

Three approaches that can be used for determining the absolute amount of the penalty include:

Approach 1: Penalties Proportional to the Perceived Value of Quality

Under this approach, a dollar value would be assigned to the Patient-Level Target to represent the value of achieving that threshold from a patient’s perspective, and then the maximum penalty for failure to achieve the Target would be based on some proportion of that dollar value.

For example, if lack of significant hip pain is assigned a value of \$5,000, then if 5% of the hip surgery procedures performed under an APM resulted in significant hip pain, the penalty would be \$5,000 for each of the

patients that had significant hip pain, which would be equivalent to a penalty of \$250 for every patient (5% x \$5,000).

Approach 2: Penalty Proportional to the Payment for Planned Services

Under this approach, the maximum amount of the penalty would be set as a percentage of the amount of payment the APM participant is receiving under Component #1.

For example, if an APM for hip surgery pays \$25,000 for the procedure, that payment might be reduced by up to 20% based on the proportion of patients who have significant hip pain. If the Target is for no patients to have pain and the penalties are proportional to the percentage of patients with significant pain, then if 5% of the patients have significant pain, the penalty would be \$250 per patient (5% x 20% x \$25,000).

Approach 3: Penalties Proportional to the Penalty or Bonus for Utilization/Spending

A third approach is to tie the penalty for quality performance to any penalties or bonuses that are paid based on Component #2 of the APM.

In the example above, if the primary goal of the APM for hip surgery was to reduce hospital readmissions for complications, and if an APM participant was scheduled to receive a \$5,000 bonus payment for having reduced readmissions by more than the Target rate for readmissions, the bonus payment could be reduced by up to 100% based on the proportion of patients who have significant hip pain after surgery. For example, if the change in care delivery that reduced readmissions also resulted in 5% of the patients having significant hip pain after surgery, the bonus payment might be reduced by \$250 (5% x 100% x \$5,000). If the provider had failed to reduce readmissions and was scheduled to pay a penalty for that, the penalty could be increased if the patients were also experiencing higher rates of pain than expected, and the penalty could be reduced if the patients' pain had decreased even though the readmission rate had not improved.

Advantages and Disadvantages of the Different Approaches

Approach #1 obviously requires reaching agreement on the dollar value of achieving a certain level of quality, and that is challenging to do. Different patients will likely assign very different values to different quality problems, and decisions would also be needed about whether and how to differentiate between quality issues that cause short-term problems for patients versus longer-term problems (e.g., temporary pain vs. chronic pain). Methodologies designed to assign relative weights to quality issues (e.g., Quality-Adjusted Life Years or QALYs) still require assigning a dollar value to a life.

However, the other approaches do not actually avoid this problem, they merely disguise it. Making the penalty proportional to the payment for the service implicitly assigns a dollar value to the quality problem based on a percentage of the payment for the service. If the same

maximum percentage penalty is used for every aspect of quality that is being measured, that implicitly means that none of the problems with quality is more serious than the others.

In addition, Approach #3 makes the value assigned to a quality problem contingent on whether the provider has deviated from the Targets on utilization and spending. If the provider's utilization and spending rates are exactly as expected, so there are no bonuses or penalties based on those rates, then there is also no penalty for problems with quality, no matter how serious those problems may be.

Example of Approach #1: In the CMS Comprehensive Primary Care Plus Initiative, participating primary care practices receive a monthly Performance Based Incentive Payment (PBIP) for each beneficiary equal to either \$2.50 or \$4.00 in addition to payments for visits, procedures, and care management services. The primary care practice is required to return up to one-half of this payment if it fails to meet performance standards on a group of measures of care quality and patient experience. This is, in effect, Approach #1: high quality of care is assigned a value of \$1.25 - \$2.00 per patient per month.¹⁸³

Example of Approach #2: In the Merit-Based Incentive Payment (MIPS) program, a physician receives a Quality Score based on their performance on a series of quality measures. If this Quality Score falls below a performance threshold (i.e., a Target Level), the fees paid to the physician for services during the coming year are reduced in proportion to the amount by which the Quality Score falls below the Performance Threshold.¹⁸⁴

Example of Approach #3: In the Medicare Shared Savings Program, Accountable Care Organizations are assigned a Quality Score from 0% to 100% based on their performance on a series of quality measures. If an ACO has achieved the Minimum Savings Rate and thereby qualifies to receive a shared savings payment, the maximum amount of the payment based on the savings achieved is multiplied by the Quality Score to determine the actual payment made. If the ACO does not achieve the Minimum Savings Rate, then it does not receive a shared savings payment, and the Quality Score is irrelevant, i.e., there is no penalty for poor quality.¹⁸⁵ A similar approach is used in the Bundled Payments for Care Improvement – Advanced APM.¹⁸⁶

d. Determining the Penalties When There Are Multiple Quality Measures

Even if the APM is focused narrowly on a particular aspect of avoidable spending, there may not be any one quality measure that adequately captures all of the important aspects of quality that are important to a patient. If multiple aspects of quality are being assessed, then a decision must also be made as to how the over-

all penalty will be determined. The three basic options are:

- i. the penalties, if any, for each aspect of quality will be determined independently and then added together;
- ii. the penalties will be determined independently and added, but the total will be capped at an overall maximum penalty amount; or
- iii. the overall penalty will be determined by evaluating all of the quality measures jointly in some way, such as through a composite measure.

Setting Overall Maximum Penalties

Without an overall maximum, the combined amount of penalties might exceed the payments the provider is receiving under the APM or even the total revenues the provider is receiving for all of the services the provider delivers to all patients, not just the patients who are participating in the APM. This could be particularly true for Approach #1; since the penalty amounts are not set in any explicit relationship to the amount the provider is being paid, even a single penalty amount for a serious quality issue might represent as much or more than the provider is being paid.

However, an overall maximum has the perverse effect of devaluing individual quality problems when multiple problems are present. If the total penalties the provider would have been eligible to pay for multiple quality problems are greater than the maximum amount, then by paying only the maximum penalty, the provider is inherently paying nothing for some of those problems, or equivalently, paying smaller penalties for each of them than other providers who performed better on the individual measures would have to pay.

Using Composite Measures

An alternative approach is to first combine the quality measures into a composite measure of quality, and then determine the penalty based on the composite measure.

Measures of *spending* on different types of services can be easily combined into a composite measure since the individual measures are all based on dollars. Even measures of utilization for different types of services can be combined into a “resource use” measure using weights defined by the relative costs or payment amounts for those services, as described in Section VI.B.

However, it is not as easy to combine measures of different aspects of quality such as mortality, pain, and mobility. Creating a composite measure requires translating all of the components into a common scale using some form of weighting system. However, if different patients evaluate the tradeoff between different aspects of quality differently, then differences in the composite measure will not accurately reflect how individual patients would rank different providers on quality.¹⁸⁷ For example, one patient may be willing to accept a certain level of pain in return for greater mobility, another may accept limited mobility in order to be free of pain. Differences in the way people evaluate pain vs. death often lead to different preferences about the choice be-

tween hospice care or continued treatment for an advanced illness.

In addition, even for quality problems that have an impact on spending, such as injuries or infections, the amounts spent to treat them may not accurately reflect the significance of the problems from the patient’s perspective. For example, a treatment-related injury that is disfiguring may have a significant negative impact beyond the cost of treating the injury. If the accountability component of the APM uses a composite measure of spending that combines measures of spending on avoidable utilization and potential complications, the individual measures will likely need to be separated so they can be evaluated separately.

e. Using Bonuses in Addition to Penalties

If the APM participant pays penalties for failure to meet a Quality Target, but receives no reward when performance is better than the Target, the participant will be penalized for random variation in quality from one period to the next. Moreover, if extra time or cost is involved in achieving better performance, there will be little financial incentive for the provider to make that investment if it receives the same payment no matter how much better its performance is. This could be addressed by using bonus payments in addition to penalties, i.e., by making an additional payment to the provider if performance on the quality measure is better than the Target.

However, it is challenging to provide bonuses in APMs for high quality care or improvements in care because, by definition, the APM is supposed to maintain or reduce spending. Although it would certainly be desirable from the patients’ perspective if a provider improved the quality of care the patients received even if no savings were achieved, paying the provider a bonus for the improved quality in that circumstance would increase spending on those patients. If a large proportion of the providers participating in the APM received bonuses, it could potentially result in higher spending for the APM as a whole. (If an improvement in quality results in a reduction in spending, then a measure of that should already have been captured as part of the accountability component for spending.)

The Merit-Based Incentive Payment System program in Medicare addresses this challenge by making the size of bonus payments for an individual physician who is eligible for a bonus contingent on the number and amounts of the penalties paid by other physicians.¹⁸⁸ If no one is penalized, then no one receives a bonus, no matter how high the quality of care they deliver. Conversely, if many physicians are penalized, then those with high levels of quality can receive large bonuses. This is a very problematic approach for several reasons:

- It discourages collaboration among physicians in finding and disseminating methods of improving healthcare for all patients. A physician who develops a better way to deliver care to patients will only be rewarded for that improvement if other physicians don’t do the same thing, since bonuses are based on whether there are differences in the quality perfor-

mance between physicians, not whether the overall level of quality has improved.

- It makes it impossible for any individual physician to predict the financial impact of developing an improved approach to care. If the physician has to incur additional costs to improve quality that are not covered by standard payments, the only way to cover the costs may be through receipt of a bonus payment, but the physician will not receive a bonus if other physicians make similar improvements. Consequently, the “safest” approach for every physician is to keep delivering care in the exact same way.

An alternative solution is to define a fixed pool of bonus funds as part of the APM. The size of the pool would be smaller than the total amount of net savings expected from the APM, so that even if quality performance was high enough to require use of all of the funds in the pool for bonuses, the total spending under the APM would still be no greater than it would have been in the absence of the APM. The amounts of bonus payments to individual APM participants could then be determined by allocating the bonus pool among all of the participants who qualified for a bonus.

Option 2: Outcome-Based Payments

Under Option 1, determining penalties based on the APM participant’s performance against a Population-Level Target for *all* patients does not result in a “value-based payment” from the perspective of an individual patient. Individual patients should expect to pay less, or nothing at all, if they receive poorer quality care than other patients, but under Option 1, some patients could receive worse care but still pay as much as patients who receive better care.

Even if a Patient-Level Target is used, if the amount of the penalty is less than the total amount the provider was paid for services, the patient for whom the Target was not achieved would still be paying for care that failed to achieve the results it was supposed to.

Under Option 2, the healthcare provider in the APM would only receive a payment under Component #1 for delivering a service supported by the APM if the Patient-Level Target was achieved for the patient who received the service. Although the option is labeled “Outcome-Based Payment,” the option can be used for a Target based on a process measure rather than an outcome, i.e., if the desired process is not performed, there would be no payment for that patient. This approach is the equivalent of a full money-back guarantee on the service (s) supported by the APM.

a. Avoiding Underpayment of Providers Without Increasing APM Spending

From the provider’s perspective, a problem with an Outcome-Based Payment is that if the provider does not expect to achieve the outcome for every patient (i.e., the Population-Level Target for the measure is not 100%), a payment under Component #1 that is based on the average cost of the services would be too low, since the provider would not receive sufficient revenues for the patients who did achieve the outcome to cover the costs of

services delivered to those patients as well as the patients who did not achieve the outcome.

This is the same issue discussed under Option 2 in Section VI.B with respect to accountability for utilization/spending. There, the solution was to increase the payment amount under Component #1 for the planned services to reflect the percentage of patients for whom the provider participating in the APM expected to meet the utilization/spending target. However, if the same approach were used for outcomes or quality, a provider that achieved the desired outcome for a higher-than-expected proportion of patients will receive a higher payment, but there would be no offsetting reduction in spending (since performance on the quality measure does not have a direct relationship with spending), and this could increase overall spending under the APM.

In addition, success on many outcome measures is not completely within a provider’s control, e.g., because the outcome depends not only on the services delivered by the provider but on actions taken by the patient. One provider could perform better on an outcome-based payment because of the types of patients they treat or the resources available to those patients, not because of the way they deliver services. In these cases, adjusting the payments based on average expected performance could inappropriately penalize some providers based on the types of patients they treat.

Example: In the Medicare Diabetes Prevention Program, a participating provider only receives payments for delivering services to a patient during the patient’s second year of participation in the program if the patient achieves or maintains a 5% reduction from their baseline weight when they entered the program. If a patient does not achieve or maintain the 5% weight loss, the provider receives no payment for the services delivered to that patient.¹⁸⁹ However, the outcome target is the same regardless of the extent of the barriers faced by individual patients in achieving the target. Providers that cannot achieve the 5% weight loss for a sufficient number of patients will not receive enough revenues to cover the cost of the services.¹⁹⁰

There are three approaches that can be used to define outcome-based payments that avoid inappropriately underpaying or penalizing providers for care while also avoiding increasing spending under the APM:

- Approach #1:** Design and use measures that can and should be achieved for every patient;
- Approach #2:** Make payments partially contingent on outcomes
- Approach #3:** Increase the patient’s cost-sharing amount for services that achieve additional quality goals

COMPARISON OF PAYMENT PENALTIES AND OUTCOME-BASED PAYMENT

The table shows a hypothetical service for which providers are paid \$200. Providers are expected to achieve a particular quality standard in return for the payment. Under Option 1, the provider is paid for every service delivered, but the amounts of the provider's payments are reduced if the quality standard is not met for all patients; the penalty is a 1% reduction in payment for each percentage point below 100% success (i.e., a Population-Level Target). Under Option 2, the provider is only paid for delivering the service to an individual patient if the quality standard is met for that individual patient (i.e., a Patient-Level Target). Under Option 1, if the provider does not meet the quality standard for 10% of the patients, the provider's revenues are reduced by 10%, but the patients whose service does not meet the quality standard pay the same amount as the patient whose service did meet the quality standard. Under Option 2, the provider's revenues are also reduced by 10%, but the patients whose service did not meet the quality standard pay nothing, and the patients whose service did meet the quality standard pay the full amount.

	Current FFS	Option 1: Payment with 10% Penalty for Poor Quality	Option 2: Outcome-Based Payment
Patients			
Total # of Patients	100	100	100
% Receiving Services Meeting Quality Standard	90%	90%	90%
# Receiving Services Meeting Quality Standard	90	90	90
Payment for Services	\$200	\$180	\$200
# Receiving Services NOT Meeting Quality Standard	10	10	10
Payment for Services	\$200	\$180	\$0
Revenue to Provider	\$20,000	\$18,000	\$18,000
% Change in Revenue		-10%	-10%

Approach #1: Design and Use Measures That Can and Should Be Achieved for Every Patient

If it is reasonable to expect that the APM participant can achieve a particular outcome or performance standard for every patient with the services and payments under the APM, then there is no need to adjust the payment amounts under Component #1. The fact that providers are not currently achieving this level of performance does not necessarily mean they could not do so in the future; if there are barriers to better performance in the fee-for-service payment system and the APM removes these barriers, then uniformly high performance could be expected. In this case, the standard payment under the APM would provide adequate revenues to cover the cost of services and the APM participant could reasonably expect to receive the payment for every patient. This is equivalent to how the provider would fare under Option 1 with a Population-Level Target set at 100% success.

However, for many current process and intermediate outcome measures, no one expects that a provider could or should achieve 100% success because the measure fails to exclude patients for whom the quality standard is inappropriate. Refining the definition of the measure would enable setting a Target based on 100% success.¹⁹¹

If success in achieving an outcome is only partly within the control of the provider, then the measure could be redefined so that it excludes patients who did not carry out actions that were essential to success. For example, if good outcomes from hip surgery depend on the patient attending rehabilitation sessions, then patients who fail to attend those sessions could be excluded from the denominator of the outcome measure. If there is strong evidence showing that a particular process or intermediate outcome has a strong causal relationship with the desired outcome, then it may be preferable to make the payment contingent on delivery of that process or achievement of that intermediate outcome rather than trying to tie the payment to the ultimate outcome.

Unfortunately, it is common practice in current quality reporting and value-based payment systems to stop using quality measures when they are “topped out,” i.e., when almost all providers are able to achieve nearly 100% success on the measures. The argument for no longer using them is that there is no longer a need for improvement. However, if the measures were originally developed because the aspect of quality was important for patients, and if there is a risk that efforts to reduce spending under the APM could negatively affect that aspect of quality, the “topped out” measures may be some of the most important quality measures to use in the context of APMs.

Approach #2: Make Payments Partially Contingent on Outcomes

The lower the expected rate of success on the measure, the bigger the financial impact on the provider if there is no payment at all when success is not achieved. This impact can be mitigated by reducing rather than eliminating the provider’s payment under Component #1 when the outcome is not achieved. For example, if the expected success rate is 90% and the provider refunds 50% of the payment when the outcome is not achieved, the financial impact on the provider would be equivalent to a 100% refund on a measure with an expected success rate of 95%. A smaller adjustment to the provider’s payment amount also reduces the potential increase in spending for a higher-performing provider. (This approach is equivalent to Option 1, except that the amount of the penalty would be defined in a way that allows a direct transition to a true outcome-based payment once there is a reason to believe that a success rate close to 100% is possible.

A partial payment (or equivalently, the partial refund of payment) would be particularly appropriate for outcomes where the provider shares responsibility for achieving the outcome with the patient. For example, a 50% refund when an outcome is not achieved might be appropriate when actions by both the provider and patient equally contribute to achieving the outcome.

Approach #3: Increase the Patient’s Cost-Sharing Amount for Services That Achieve Additional Quality Goals

If there is a quality goal that the provider cannot reasonably expect to achieve for every patient at the standard amount of payment, but achieving that goal is important to some patients, then the provider could be paid more for the services when the outcome is achieved, but the increase in payment would be reflected in higher *cost-sharing by the patient*, rather than an increase in the *payment by the patient’s insurance plan*. In this situation, if the provider did not achieve the quality goal, the health plan would still pay for the service (assuming that the service met other quality standards), but if the service achieved the quality goal, the provider would receive a higher payment, and the difference in payment (which would be equivalent to a bonus payment) would be paid by the patient, not by the insurance plan.

Example: If standard hip replacement surgery can routinely be expected to enable patients to walk without pain, but not to run long distances, and if there is a more expensive form of the surgery that can enable a patient to run (e.g., using a different type of prosthetic hip), a patient who wanted to be able to run could pay the extra cost of the surgery contingent on achieving the higher-level outcome. The surgery provider would receive no payment at all if the patient was unable to walk, the provider would receive the standard payment from the patient’s insurance plan if the patient could walk, and the provider would receive an additional payment from the patient if the patient could both run and walk.

This approach would reflect the fact that the aspect of quality being measured does not have a direct impact on the healthcare spending supported by the insurance plan. (If it did impact healthcare spending, then this aspect of quality should presumably have been included in the utilization/spending measures under Component #2.) The benefit of higher quality would accrue entirely to the patient, not to their insurance plan, so the patient would be agreeing to pay more themselves in order to receive care from a provider who would only charge the patient if that aspect of quality was achieved.

In addition, this approach would provide a method of determining the relative value that patients assign to different aspects of quality. If a patient is unwilling to pay more for a service with a money-back guarantee for a particular aspect of quality, then that aspect of quality may not be important to the patient.

This approach could not be used for patients who would otherwise pay nothing for the healthcare services (e.g., patients on Medicaid who have no cost-sharing obligation). However, as long as a significant portion of a provider’s patients were willing to pay in this way, the provider would still receive a bonus for delivering higher-quality care to everyone.

b. Partial/Interim Payments for Longer-Term Outcomes

In some cases, it will take months or even years to know whether a patient achieved all of the desired outcomes from a treatment. For example, although post-surgical infections will generally appear fairly soon after hip or knee surgery is completed, and a patient may be able to walk again in weeks, days, or even hours after the surgery is finished, it may take months before a person is able to perform all of the activities they want or need to participate in. Cancer patients want to survive for many years after treatment, not just for a few months.

However, most healthcare providers cannot wait for months or years in order to be paid for their services. This could be addressed by making a partial payment for treatment when short-term outcomes are achieved and an additional payment when a longer-term outcome is achieved, with the amount of the additional payment set so that it incorporates the interest lost or paid during the interim period.

A variation on this approach would be to distribute the additional payment for the longer-term outcome on an amortized basis each month in which that outcome is achieved, e.g., the Team would receive an additional payment for each month in which the patient is alive after treatment for a high-mortality illness. This is similar to an approach that Rolls Royce has successfully used in its jet aircraft business. Rather than charging a fixed price for an engine, an airline pays based on the number of hours the engine successfully operates. The longer the engine functions without problems, and the sooner that it is returned to service when it does have a problem, the more money Rolls Royce makes.¹⁹²

c. Limiting the Provider's Losses

Under Option 1, it is relatively easy to limit the total financial risk for an APM participant, because the APM participant pays a single aggregate penalty to a payer at the end of a performance period, and that penalty can be limited to a maximum amount. Under Option 2, however, any penalty is determined on a patient-by-patient basis, so as more patients fail to achieve the desired outcome, the loss of revenue would continue to grow.

Limiting the total financial risk under Option 2 requires use of a stop-loss insurance mechanism. Under stop-loss insurance, if the provider participating in the APM reaches a certain threshold of losses, the participant receives additional funds from an insurer to cover all or part of the losses above that threshold. If the stop-loss protection is not provided by the patient's health plan, the premium that the APM participant pays to a separate insurer for the stop-loss insurance would then need to be factored into the cost of delivering services under the APM.¹⁹³

Option 3: Warrantied Payments

The Outcome-Based Payment approach in Option 2 assures each individual patient and their health insurer that if the new/revised services delivered to the patient through the APM fail to achieve the outcome or standard of quality, the patient/payer will not have to pay for the new/revised service(s). However, the patient would still fail to achieve the outcome or experience the quality problem. Although this may not have a direct or immediate effect on healthcare spending (otherwise it should have been included as part of the utilization/spending measures for the APM under Component #2), it could have other negative impacts on the patient. Merely eliminating the payment for the services delivered by the provider does not create a neutral result for the patient.

This can be addressed by defining a "warrantied" payment that requires the APM participant to not only deliver the desirable services in return for the payment under Component #1, but also to pay some type of compensation if the desired outcome is not achieved or poor-quality care is delivered. In Section VI.B, the warranty was based on paying for the costs of healthcare services needed to treat problems; for a quality problem that is not treatable, an amount of payment would have to be assigned to the quality problem using approaches similar to those discussed under Option 1 for determining penalties. Option 3 differs from Option 1 in that the

penalties would be determined and paid on a patient-by-patient basis rather than based on the provider's aggregate performance for all patients.

Under the warranty payment approach, the provider might still receive the standard payment under Component #1 for the services that were delivered to the patient, but the provider would pay compensation to the patient at the point when it was clear that the outcome was not achieved or a quality problem had occurred. Depending on the size of the compensation payment required when the quality problem occurs, the financial impact on the provider could be more or less than the outcome-based payment approach in Option 2.

Option 3 (the warranty payment approach) could also be combined with Option 2 (the outcome-based payment approach), so that when the desired outcome is not achieved, the provider receives no payment for the services delivered and also pays the penalty specified in the warranty. This would be particularly appropriate when the patient will want or need to pursue treatment from another provider, since not only will similar services need to be purchased again, but additional services or more expensive services may be needed to correct the problems created during the initial, unsuccessful treatment.

Limiting the Provider's Losses

Similar to the discussion under Option 2, it would be preferable to tie the warranty payments to quality or outcome measures that the provider could expect to achieve for every patient, so that there would be no need to adjust the basic payment amount under Component #1 to reflect both the cost of the desirable services and the expected amount of warranty payments for quality problems. However, as long as there is the possibility of quality problems, there would be the possibility of warranty payouts, and the APM participant would need to have a way of paying for them when they occur. This could be addressed by enabling the APM participant to purchase stop-loss insurance. Under stop-loss insurance, if the APM participant reaches a certain threshold of losses, the participant receives additional funds from an insurer to cover all or part of the losses above that threshold. The premium that the APM participant pays for the stop-loss insurance would then need to be factored into the cost of delivering services under the APM.

Similar to the discussion in Section VI.B, two different types of stop-loss insurance might be needed:

1. **Individual stop-loss** insurance, to address large warranty payments for individual patients who experience major quality problems; and/or
2. **Aggregate stop loss** insurance, to address large total warranty payments when multiple patients experience quality problems.

As with Option 3 for Component #2, the provider could be viewed as offering "insurance" by at least some, if not all, state insurance commissioners if the warranty obligates the provider to pay large amounts when a quality problem occurs and the threshold for stop-loss insurance is high.

Option 4: Termination of the Provider's Participation in the APM

A fourth option is to terminate participation in the APM by a provider or group of providers if the provider does not achieve success on one or more of the quality performance measures. As noted in Section VI.B, this is essentially the approach that currently exists in the standard fee-for-service payment systems used by Medicare and private payers. A provider must meet some set of minimum standards of performance (e.g., hospitals must be accredited) in order to continue being paid for services. For many aspects of performance that are assessed through this process, there is no absolute threshold of performance that must be achieved, but instead, if problems exist, a determination is made as to whether the problems are remediable and are being remedied in order to decide whether termination should occur. Termination can represent a significant financial penalty if the provider had hired new staff, purchased new equipment, or incurred other kinds of costs that were not fully covered by the payments received prior to termination.

An advantage of Option 4 is that it allows greater flexibility to consider the circumstances that may have led to failure or success in meeting the Quality Targets. For example, if a provider was delivering services to patients who faced unique challenges in achieving standard levels of outcomes, and the risk adjustment or stratification system in the APM did not adequately address these challenges, Option 4 would enable a determination to be made that no penalty was justified when the provider's performance fell short of the Target. Conversely, if it turns out the provider achieved the Target by systematically avoiding patients who were likely to face greater challenges in achieving the desired outcome, the APM participant could be terminated even though a standard formula might have determined that performance was satisfactory.

Option 4 may be the only feasible option for small providers, for APMs focused on small numbers of patients or health conditions that occur relatively rarely, or for APMs designed to reduce problematic outcomes that occur rarely or unpredictably. No matter how sophisticated the statistical methodology, it may be impossible to create a fair way of determining penalties or basing payments on outcomes, and so a different form of evaluation may be necessary. For example, outcomes might be evaluated over a multi-year period, or detailed clinical audits of individual patients might be used to verify that the most appropriate care was being delivered.

Option 4 could also be used during the initial years of implementation of an APM, or during the type of "beta testing" process described in Section VIII.B, when there is uncertainty about how changes in care delivery will affect quality. A transition could then be made to one of the other Options.

Option 5: Termination or Modification of the APM

A final option is to stop the use of the APM altogether. If the APM saves money but the quality of care delivered by most participating providers has worsened, or if the APM fails to reduce spending or improve quality, then it would not meet the criteria for an APM. It might be possible to correct the problems by making significant revisions in the design of the APM, or an entirely different approach may be needed.

Option 5 may also be the only practical mechanism for dealing with APMs where the primary outcomes are long-term in nature. If care must be delivered in a different way over an extended period of time in order to affect an outcome, or if the difference in outcome will not manifest itself until several years in the future, there will be no way to adjust a provider's current payments based on performance nor to distinguish which providers are under-performing until the outcomes can actually be measured. When the outcome data are available, a decision can be made as to whether the APM should be terminated or modified.

TABLE 12
METHODS OF PERFORMANCE-BASED PAYMENT FOR QUALITY

Alternative	Strengths	Weaknesses
1a. Penalties Based on Failure to Meet Targets	The amount of the penalty can be made small or large based on factors such as the likelihood of achieving the Target and the ability of the provider to afford penalties	Some patients and payers will have to pay for low-quality care even if a Population-Level Target is achieved There is no reward for performance that is better than the Target
1b. Penalties & Bonuses Based on Achieving Targets	Encourages generating better outcomes than the Target level	Bonuses for high-performing APM providers could potentially increase overall spending
2. Outcome-Based Payment	Ensures there is no payment (or a large reduction in payment) for an individual patient if the Target is not achieved for that patient	Works best when there is a high probability of achieving the Target for most patients Does not compensate the patient or payers for costs they incur for treating complications and other costs not directly supported by the payment
3. Bundled/Warranted Payment	Compensates the patient for the provider's failure to achieve a promised outcome rather than just refunding payment for ineffective services	Could create significant financial risk for providers if the compensation for failure to achieve outcomes is large or occurs frequently
4. Termination of Provider's Participation in the APM	Allows the determination of whether outcomes were achieved to be based on a more detailed evaluation than is possible through comparing quality measures to a Target	Could result in patients being harmed before a decision is made to terminate a provider Premature termination could result in a lost opportunity to achieve better outcomes in the future
5. Termination or Modification of the APM	Avoids continued harms to patients from an APM that fails to maintain or improve quality of care or outcomes	Providers could be less likely to participate in the APM or to transform care delivery significantly if it is uncertain whether the APM will continue to be offered in the future or if major changes will be made

D. APM Component #4: Defining the Eligible Patients

There is tremendous diversity among patients in terms of both the types of health problems they have and other characteristics that affect the cost of treatment and the outcomes that can be achieved. The broader the range of patients that an APM attempts to include, the more complex the structure of the APM will likely need to be. In general, the determination of which patients should be eligible or ineligible for participation in an APM should be made after at least a preliminary design of the other components has been completed, so that the tradeoffs between breadth of patient eligibility and the complexity of the APM structure can be assessed.

1. Establishing Appropriate Eligibility Criteria

A well-designed APM provides an opportunity to reduce unnecessary services and spending without harming patients. However, even if the APM is successful in reducing use of unnecessary services for many patients, there is the risk that the service(s) supported by the APM will be overutilized in ways that can lead to higher spending or worse outcomes. There will inherently be variation across patients in the benefits the APM will provide versus the costs the patients or their payers will incur to participate, so the greater the participation by patients who achieve lower benefits at higher costs, the smaller the overall benefits the APM will produce, and the greater the possibility that on average, the APM will reduce rather than improve value. In order to address this, eligibility criteria could be defined that limit participation to the patients who would have been most likely to receive the unnecessary services and/or to benefit from the services supported by the APM.

Example: Suppose that an APM would enable a primary care practice to deliver home-based services to help patients with a chronic disease to avoid hospitalization. An analysis of the primary care practice's past patients showed there was a clear business case for paying for these services, i.e., the savings from lower hospitalizations would more than offset the higher cost of the home-based services. However, if patients who are not at high risk of hospitalization begin seeking care from the primary care practice simply because it now offers the home care services, the number of patients participating in the APM will increase, and the average rate of hospitalizations will decrease. If this increase in participation is large enough, the APM could appear to be effective (because the rate of hospitalizations for the practice's patients would be lower than in the past due to the change in mix of patients), even though it is actually increasing total spending (because it is providing an additional service to many patients who would not have been hospitalized anyway). This could be addressed by limiting eligibility for the APM to patients with more severe chronic diseases or other characteristics that place them at higher risk of hospitalization.

a. Basing Eligibility on Patient Diagnoses and Needs Instead of Services Delivered

In traditional fee-for-service payment, eligibility for payment is based on the delivery of a specific service. However, if one of the goals of the APM is to avoid overuse of a service, or to provide flexibility to deliver different types of services or to deliver a service in different ways, eligibility for the APM will need to be defined in terms of something other than delivery of a specific service. For example, eligibility could be based on whether the patient has a specific health problem or combination of problems or whether the patient has other characteristics that would justify the use of the APM.

In most cases, it is unlikely that the mere presence or absence of a particular disease or health problem will be sufficient to serve as the eligibility criterion for an APM. A patient with a particular disease or condition may have a very mild form or a very severe form of that disease/condition, and the services they need and the outcomes that can be achieved will generally depend on the severity of the condition. The opportunities for savings or quality improvement that prompted the development of the APM, and the business case for a different payment and care delivery model, may only exist for patients with particular combinations of characteristics, and so eligibility to participate in the APM may need to be limited to those patients in order to avoid overuse of the APM.

This is conceptually similar to the criteria currently used to determine whether a patient is eligible for payment for inpatient hospital care. A diagnosis alone is generally not sufficient; the hospital must verify that the severity of the disease and other characteristics of the patient meet criteria justifying treatment in an inpatient setting rather than in an outpatient setting or in their home.

b. Avoiding Overdiagnosis and Overtreatment Due to Narrow Eligibility Criteria

Although the goal of establishing minimum eligibility criteria is to avoid overuse of the APM, caution is needed to ensure the criteria themselves do not encourage overdiagnosis or overtreatment. Narrowly-defined eligibility criteria can create a significant "cliff" in benefits – i.e., patients who meet the criteria can receive improved services under the APM whereas patients who fall just short of the criteria receive nothing. This can create a perverse incentive for both the patient and the provider to find ways for the patient to meet the criteria and participate, which in turn can result in higher spending than might otherwise occur.

Example: Patients with chronic diseases such as heart failure or COPD are eligible to participate in the CMS Bundled Payments for Care Improvement - Advanced APM, but only after they are admitted to the hospital for the condition and only for a limited period of time after discharge.¹⁹⁴ Patients with chronic disease who have been hospitalized have a significantly higher risk of a subsequent hospitalization than patients who have never been hospitalized, and

so an APM that provides the flexibility to deliver additional and different services after discharge could reduce readmissions and improve outcomes for patients. However, since patients who have not been hospitalized can also benefit from such services, limiting the services to patients who have been hospitalized creates a perverse incentive to hospitalize a patient at least once so that they can qualify for the services supported by the APM.

Defining eligibility based solely on the patient's health problems (a "condition-based" APM rather than a "procedure-based" APM) avoids the requirement that a patient receive a particular procedure or service (e.g., being hospitalized at least once) in order to qualify for the enhanced services under the APM. However, if the diagnosis of a health condition is based in part on subjective professional judgment rather than objective, independently-verifiable test results, there is the potential that a clinician will begin assigning the qualifying diagnosis to borderline cases, thereby increasing the number of patients who are eligible. As discussed in Section VI.C, the quality accountability component of the APM may need to include methods for assuring the accuracy of information about diagnoses and other patient characteristics.

c. Payment Stratification Versus Eligibility Criteria

The more restrictive the eligibility criteria, the bigger the "cliff" that is created for a patient who falls just short of the eligibility criteria and the greater the incentive for providers to manipulate the criteria in some way in order to enable a patient to qualify for a service that could benefit them. An alternative to narrowing the eligibility criteria for participation in the APM is to stratify the payment amounts and accountability measures in the APM, using the approaches described in Sections VI.A, VI.B, and VI.C. Separate tiers could be defined for patients who are less likely and more likely to have high levels of avoidable spending. The payment amounts for the patients in the lower-risk tiers could be smaller and the benchmark for avoidable spending could also be lower, so that there is a net benefit regardless of how many patients are in each tier.

Payment stratification does not eliminate the challenge of defining appropriate eligibility criteria; in fact, it complicates it because multiple eligibility criteria must now be defined, one for each payment tier. However, this is a more realistic reflection of the differences in patient needs than the binary in/out decision required by a single eligibility criterion. Moreover, while payment stratification can reduce the problem of eligibility cliffs, it cannot eliminate them. If the services and accountability standards in the lowest-risk tier are not significantly different from patients who are not eligible at all, the patients in the lowest-risk tier could be misled into believing they will receive something more than they will, and the administrative costs associated with making the payments for the patients and tracking spending and quality measures in that tier may exceed the benefits it achieves.

d. Why Population-Based Payments Don't Avoid the Challenges

"Population-based payments" might appear to avoid these challenges altogether by simply making one fixed payment for every patient in order to cover the costs of all of the services that the entire group of patient needs, rather than trying to define specific groups of patients with particular conditions or other characteristics and making separate payments for different kinds of services for patients in each group. However, global capitation and similar population-based payments do not avoid the challenges described above, they simply push them down one level – the provider or group of providers that receives the population-based payments still has to decide how to allocate those payments among all of the providers who are delivering services, and that inherently includes making decisions about which patients should be able to receive which services. If a population-based payment is only intended to cover a subset of services, then the same issues described above still exist – a decision has to be made about the eligibility criteria for eligibility for patients as well as decisions about the kinds of services the payment is expected to support.

2. Prospective vs. Retrospective Eligibility Determination

In many current Alternative Payment Models, the determination of whether a patient is participating in the APM is made retrospectively, i.e., after services have already been delivered. For example:

- In the Medicare Shared Savings Program (MSSP) APM, CMS "attributes" a Medicare beneficiary to a primary care practice if the beneficiary made more visits to that practice during the year than to any other primary care provider. The Accountable Care Organization (ACO) is then held accountable for spending on all types of services these beneficiaries received during the year and for various aspects of the quality of care for the attributed beneficiaries during the year.¹⁹⁵ A similar process is used in the Comprehensive Primary Care Plus (CPC+) initiative.¹⁹⁶ Multiple studies have shown that these attribution formulas have a high rate of errors.¹⁹⁷
- In the CMS Bundled Payments for Care Improvement – Advanced APM, patients are only eligible if they are hospitalized for procedures or conditions in one of a group of eligible Diagnosis Related Groups (MS-DRGs).¹⁹⁸ However, patients are not assigned to an MS-DRG until after they are discharged from the hospital, and the assignment is performed by "grouper" software that considers all of the procedures performed in the hospital and the diagnosis codes for all of the health problems the patient had during their hospital stay. As a result, the hospital and physicians participating in the program cannot accurately determine whether a patient will be eligible for the bundled payment until after the patient has already been discharged from the hospital.

The primary rationale for using these approaches is that they can be implemented by a payer without the need to

It is important to recognize that the “prospective attribution” methodologies that are being used by CMS and other payers only provide a partial solution to the problems of retrospective attribution methodologies described above and they introduce new types of problems. Under typical prospective attribution methodologies, a provider is told at the beginning of a performance period which patients are being attributed to them based on the *kinds of services the patient received during the previous year*. Although knowing who your patients are at the beginning of the performance period is better than not finding out until the end of the period, the methodology is still based on a statistical formula and claims data that can make errors in the eligibility determination. Moreover, because the attribution determination is based on data about what happened to the patient in the past, rather than during the period in which the APM services were being delivered, additional errors will be made for patients whose characteristics have changed.

Section VII describes how true prospective eligibility determinations can be operationalized in ways that avoid these kinds of problems while also reducing administrative burdens for both payers and providers.

3. Avoiding Cherry-Picking and Lemon-Dropping

Although carefully designed eligibility criteria in an APM can help avoid the APM being used for too *many* patients, they do not protect against the APM being used for too *few* patients. With the exception of hospital emergency departments governed by EMTALA (the Emergency Medical Treatment and Labor Act)²⁰¹, healthcare providers are not obligated to deliver services to every patient who might need or want them, so there is the possibility that some providers participating in an APM could selectively avoid patients who need more services and/or are less likely to have favorable outcomes (this is sometimes referred to as “lemon-dropping”), and to focus their services on the patients who are likely to have the most favorable outcomes (i.e., “cherry-picking”). The provider who engages in cherry-picking could receive smaller penalties or larger bonuses than a provider who does not, which in turn could cause the payer to spend more than they would have otherwise. In addition, the patients with unfavorable characteristics may have greater difficulty obtaining the care they need.

This issue arises with virtually any payment system, not just Alternative Payment Models. For example:

- In the fee-for-service system for physicians, different amounts of time will be needed to deliver the “same” service to different patients, so the physician will benefit if he or she can choose to deliver the service only to those patients requiring the least amount of time. There is no adjustment for the additional time a physician may need to spend with a patient who does not speak English or who has significant social needs, so physicians who treat such patients are penalized financially for doing so.
- In the prospective payment system for inpatient hospital care, the hospital receives the same amount of

payment for patients who are assigned to the same DRG, even though patients in the same DRG may have very different lengths of stay and require more or less intensive care.

- In any pay-for-performance system that creates accountability for cost and/or quality, there is a financial incentive for a provider to selectively choose patients who need fewer-than-average services or who have a better-than-average likelihood of achieving good outcomes for a given level of services.

One of the purported advantages of retrospective attribution and eligibility determination systems is that providers cannot selectively enroll desirable patients. However, since the retrospective attribution system only assigns a patient to a provider if that provider delivered at least one service to the patient, a provider can easily avoid having a patient attributed to them by never delivering a service to that patient (or by not billing for a service that is delivered). This could actually reduce access for patients, since it eliminates the ability for a provider to deliver a specific service to a patient who needs only that one service without having to enroll the patient in the overall program of care supported by the APM.

The right way to avoid cherry-picking and lemon-dropping is to (1) identify the factors that affect how many services a patient will need and the outcomes they will experience and (2) incorporate those factors into the design of the APM, so providers receive appropriate payments for higher-need and lower-need patients:

- If patients with specific characteristics will require more time or resources from the provider delivering the desirable services, the provider should receive a higher payment for those patients under Component #1. Section VI.A describes how payment amounts can be aligned with the costs of delivering services.
- If patients with specific characteristics have a higher risk of complications or other outcomes that increase total spending or if it is harder to achieve quality-of-care goals for them, then the performance Targets under Component #2 and/or Component #3 should be adjusted to reflect those differences, as described in Sections VI.B and VI.C.

If payments and performance measures are effectively stratified or risk-adjusted based on the patient characteristics that have a significant effect on costs and outcomes, then there would be no financial advantage to the provider from excluding patients based on those characteristics. Indeed, the provider could actually benefit from serving the higher-cost, higher-risk patients if they offer greater potential for performance bonuses than other patients.

It is certainly possible that some providers will be able to identify patient characteristics that have a significant effect on cost or quality but are not incorporated into the APM stratification system, and if so, they could decide to selectively focus on patients who don’t have those characteristics and financially benefit from that. It is also possible that providers who serve a disproportionate number of patients with the unfavorable characteristics would be financially harmed. In order to address this:

obtain any additional information from either providers or patients beyond what is currently recorded in standard billing and claims forms. However, this administrative advantage is outweighed by at least four serious problems:

- **Attribution methodologies make many errors in assigning patients to providers.** Attribution methodologies are based on certain assumptions and formulas about how patients receive services, and not every patient receives services in ways consistent with those assumptions and formulas.

Example: Most attribution methodologies attribute patients to the primary care provider that had the most visits with the patient during the course of a year. However, the fact that a patient made more visits to a particular primary care provider than any other provider during the course of an entire year does not mean that primary care provider was managing the patient's care during the entire year or that the primary care provider was managing the patient's care during the period of time when the patient had the health problem that the APM is focused on.

Example: In the CMS Oncology Care Model, participating oncology practices can bill Medicare for six monthly payments (called the Monthly Enhanced Oncology Services, or MEOS, payments) to support enhanced services for a patient who is receiving chemotherapy from that practice. However, if CMS determines that a patient did not receive the majority of cancer-related physician office visits from the oncology practice during the six-month period, the oncology practice is required to return the MEOS payments for that patient to CMS, even though the practice has already delivered services to the patient. Under the retrospective attribution methodology CMS uses, a 6-month episode of chemotherapy is attributed to the provider entity that delivered the most Evaluation & Management (E&M) Services visits to the patient during the 6-month episode that had a cancer diagnosis indicated on the claim form. However, patients may be receiving many services from an oncology practice that are not billable as E&M Services, and a patient with comorbidities may be receiving E&M Services from other physician practices for those comorbidities in addition to their chemotherapy, so the fact that a patient had a majority of E&M Services visits from practices other than the oncology practice does not mean the oncology practice was not managing their cancer care.¹⁹⁹

- **It is impossible for a provider to determine which patients can receive different services if they don't know which patients are eligible for the APM until after services are delivered.** If there are barriers in the current payment system to delivering the services patients need and the APM has been designed to remove those barriers, then a provider's willingness and ability to deliver the services will depend on whether the patient is eligible for the APM. If the provider won't know if the patient is eligible until after the service has already been delivered, the provider

would have to incur the cost of delivering the service without any assurance of payment, and that is likely to discourage providers from delivering such services, thereby making the APM itself appear less effective. Moreover, these attribution systems can create significant administrative burdens for providers as they try to match attribution lists to their own patient records and spend time asking payers to make corrections for inaccurately attributed patients.

Example: In the Medicare Bundled Payment for Care Improvement (BPCI) APM, CMS waived the requirement that a beneficiary have a 3-day inpatient hospital stay in order to qualify for a Medicare payment in a skilled nursing facility, thereby allowing patients to be discharged from the hospital to a SNF sooner than otherwise. However, the waiver only applied to patients who had an eligible stay in the hospital, and since eligibility was not officially determined until after the patient was discharged from the hospital and a Diagnosis Related Group (DRG) was assigned to the hospital stay, a BPCI provider had no assurance that Medicare would pay for SNF services after a less-than-three-day hospital stay. As a result, an evaluation of the program showed that the waiver was not widely used.²⁰⁰

- **Basing eligibility on the services delivered discourages innovation in service delivery under the APM.** For example, if a primary care practice can effectively manage a patient's care using a combination of telephone/email contacts with the physician and home visits by a nurse rather than traditional office visits with the physician, the patient would no longer be attributed to the practice because the attribution methodology is based on office visits.
- **The services under the APM may be specifically intended to change the characteristics of patients that are being used to determine eligibility.** For example, if poor care delivered in a hospital results in serious complications for a patient during the hospital stay, the diagnoses for those complications and the procedures performed to treat them could cause the patient to be "grouped" into a different DRG than the DRG that would have been assigned based on the initial reason for admission. As a result, in an APM focused on specific DRGs, the patients who have the most serious quality problems can be excluded from the spending and quality measures used in an APM based.

These problems are caused by a fundamental conflict between the retrospective eligibility methodology and the goal for creating an APM. Consequently, it is essential that the determination of whether a patient is eligible for an APM be made prospectively, i.e., before the provider participating in the APM begins delivering services supported by the APM to the patient. In most cases, this determination will need to be made by the provider(s) of services, not by the payer, particularly if the eligibility criteria are based on patient characteristics that are not currently recorded on standard claims forms.

- Research could be periodically conducted to determine whether there are patient characteristics that are causing variations in financial performance for providers, and if so, the APM could be revised to incorporate those characteristics into the payment and performance stratification structure. For example, research on social determinants of health has led to widespread recognition that pay-for-performance systems can unfairly penalize safety-net providers because of failure to adjust for social and economic barriers faced by patients.
- Patients could be periodically surveyed to determine if they are having any difficulties accessing care.

4. Patient Agreements as a Component of Eligibility Criteria

Traditionally, decisions about whether a particular service is eligible for payment have been made by payers. In some cases, these decisions are predetermined based on coverage provisions in the insurance plan the patient has selected or been assigned, while in other cases, the payer may use a prior authorization process to make an individualized determination for each patient. The provider's role is limited to determining whether a particular patient is appropriate for one or more of the services that the payer is willing to pay for, and potentially advocating with the payer to make a different decision about whether a service qualifies for coverage or authorization. The patient may only be involved in the decision-making process if the provider feels there are two or more appropriate services for the patient and both are eligible for payment, or if the payer has denied payment for a service the provider feels should be delivered.

Alternative Payment Models introduce a new decision about a new choice – whether a particular patient should receive services under the APM or under the traditional payment system. To date, most APMs have followed the traditional approach to making this new decision – the payer decides which patients are eligible to participate in the APM, the provider determines whether the patient is appropriate for particular services that are eligible for payment under the APM, and the patient may only be involved if the provider feels there are choices about the services that can be delivered. Under many APMs, because of the retrospective approach to eligibility determination described above, the patient has no choice about whether to participate in the APM and the patient may not even be informed about potential impacts that APM participation may have on the type of services they can receive.

However, there are important reasons why a different decision-making process – involving patients at the beginning rather than the end – should be used in APMs:

- **There are risks of undertreatment in APMs that do not exist in fee-for-service.** If the provider participating in the APM will be financially penalized when a particular service is used too frequently or when average spending on patients' care is too high, there is the possibility that the provider will withhold a service that the patient needs. If the provider is no longer being paid based on the number of services deliv-

ered, the provider could benefit financially by delivering fewer services.

- **Good outcomes depend on patients, not just healthcare providers.** In most cases, achieving a good health outcome for a patient depends on actions taken by the patient as well as the healthcare providers who are delivering services. If a patient fails to use prescribed treatments or avoid situations that exacerbate their condition, the provider could be penalized through no fault of their own.

Both of these elements can be addressed by requiring a Patient-Provider Care Agreement as part of the eligibility criteria for the APM. The patient would only be eligible for the APM (and the provider would only be held accountable for the patient's care under the APM) if the patient and provider have both signed a Care Agreement that describes:

- **Provider Commitments:** The Care Agreement could describe the types of services that the patient can expect to receive from the provider (or a team of providers) who is accountable for cost and quality under the APM, and the ways in which decisions will be made when choices about services are available. For example, if the APM provider is being paid based on the patient's condition rather than on the specific procedures used to treat the condition, the Care Agreement could require that a shared decision-making process be used to ensure the patient understands and agrees that the health condition exists and that a potentially avoidable service will not be performed to treat it.
- **Patient Commitments:** The Care Agreement could describe the specific actions that the patient agrees to take in order to increase the success of treatment. For example, the Care Agreement could require that the patient take all prescribed medications and notify the physician if any dosages have been missed or if the patient is experiencing difficulties obtaining or taking the medications.

Example: When the Geisinger Health System implemented its "ProvenCare®" program that provided warranties for various types of care delivery, a "patient compact" (signed by both the patient and Geisinger) was used to explain the actions both Geisinger and the patient or family would need to take in order to achieve a successful outcome.²⁰²

Example: In order for a physician or other clinician to receive payment for Chronic Care Management (CCM) services from Medicare, they must obtain patient consent. Consent may be verbal or written but must be documented in the medical record, and includes informing the patient about the availability of CCM services and applicable cost-sharing, the fact that only one practitioner can furnish and be paid for CCM services during a calendar month, and informing the patient of their right to stop CCM services at any time (effective at the end of the calendar month). Informed patient consent need only be obtained once prior to furnishing CCM, or if the patient chooses to change the practitioner who will furnish and bill CCM.²⁰³

E. Finalizing the APM Design

1. Combining the Components of the APM

As the previous sections make clear, there are multiple options for designing each of the four components of an Alternative Payment Model. Although not every option will be appropriate in every situation, there will generally be multiple options that could potentially be used. The advantages and disadvantages of the different options will depend on the specific types of opportunities for savings and improvements in quality that are being pursued, the approaches to care delivery that will be used to address those opportunities, and the specific barriers in the current payment system that need to be corrected.

The choice of options within each component will also depend on which options are chosen for other components. For example:

- If Component #1 addresses lack of payment for a particular service by creating a new fee for that service, Component #2 should be designed to assure that the additional spending on that service is offset by savings on other planned or unplanned services.
- In contrast, if Component #1 replaces current fees for other services with a bundled payment that can be used for the new service as well as the services for which fees currently exist, then Component #3 should be designed to assure that patients are not receiving fewer of the current services than they need.
- The more aspects of services and spending for which Component #2 holds a provider accountable, and the larger the expected reduction in spending, the more aspects of quality that Component #3 will need to ensure are maintained.
- The more types of patients that are eligible to participate under the criteria defined in Component #4, the more likely it will be that the payments in Component #1 and the accountability measures in Components #2 and #3 will need to be stratified to properly address differences in those patients' needs and the outcomes that can be achieved for them.

In most situations, there will be multiple ways to combine different options for the various components into an APM, and there will be no one best combination of options to choose. In practice, the choice of options will often depend on additional considerations, such as the operational and implementation issues described in Sections VII and VIII.

a. Components Used in Medicare APMs

The APMs implemented to date in the Medicare program have used a narrow range of the options available for the four Components.

Use of Component #1 in Medicare APMs

Most of the current APMs in the Medicare program do little or nothing to directly address the issues with current payment systems that create barriers to delivering better care:

- The largest of the Medicare APMs – Track 1 of the Medicare Shared Savings Program – makes no changes at all in the underlying payment systems for services. It simply awards shared savings bonuses or imposes financial penalties based on Component #2.
- Only two of the Medicare APMs – Comprehensive Primary Care Plus and the Oncology Care Model – provide large, flexible payments to participating providers that can be used for services that are not paid for, or not paid for adequately, under current payment systems. However, it is unclear how the payment amounts were chosen and whether they are adequate to cover the costs of the services needed to achieve the goals of the APMs.

Use of Component #2 in Medicare APMs

Table 13 shows that in most of the major Medicare APMs (the Medicare Shared Savings Program, Bundled Payments for Care Improvement – Advanced, the Oncology Care Model, and the Comprehensive Care for Joint Replacement program), the accountability measures are defined as the total Medicare spending for the patient, either for an entire year or for an “episode” lasting several months. This approach has many weaknesses, as discussed at length in Section VI.B.

The only major Medicare APM that bases accountability on utilization of specific services rather than the total cost of care is the Comprehensive Primary Care Plus (CPC+) APM. This was a change from the original Comprehensive Primary Care Initiative (CPCI) where a portion of the payments to practices were expected to come from shared savings on the total cost of care for the patients. When CMS terminated CPCI and replaced it with CPC+, it explicitly chose not use a total cost of care shared savings component, stating that shared savings was not a desirable way to pay primary care practices.²⁰⁴

Also, the potential penalties in most of the APMs are very large – if spending is higher than Target levels, a participating provider could have to pay CMS as much as 20% of the total amount Medicare is spending on services for all of the participating patients' needs. The only exception is the Comprehensive Primary Care Plus APM, where the penalty is limited to returning one-half of the Performance-Based Incentive Payment that CMS has paid to the primary care practice in advance.

**TABLE 13
ACCOUNTABILITY FOR UTILIZATION/SPENDING IN CURRENT CMMI APMS**

APM	Measure(s)	Measure Type	Target Type	Benchmark	Target Change	Target	Performance Assessment	Type of Payment Adjustment	Amount of Penalty	Limit on Penalty
Bundled Payments for Care Improvement - Advanced (BPCI-A)	(1) Episode Spending on eligible hospital admission/outpatient procedure + Total spending (except for Part D drugs and specified exclusions) during 90 days after procedure or discharge, capped at 99th percentile	Population-Level Measure	Benchmark-Based Target	Prior Performance Benchmark: Historical spending by hospital and physician on similar episodes adjusted for (1) changes in prices, (2) changes in patient comorbidities and other characteristics, and (3) peer group trend	Desired Change: 3% reduction in spending compared to benchmark	Benchmark + Target Change	Comparison of actual spending to Target, aggregated across all episode types	Penalty for spending in excess of Target Bonus for spending below Target	100% of spending above Target	Penalty limited to 20% of Target
	(2) Episode Spending Total spending (except for Part D drugs and specified exclusions) during 91st-120th days after procedure or hospital discharge	Population-Level Measure	Benchmark-Based Target	Prior Performance Benchmark: Historical spending by hospital and physician on similar episodes adjusted for (1) changes in prices, (2) changes in patient comorbidities and other characteristics, and (3) peer group trend	Statistically Significant Change: within 99.5% confidence interval of zero	Benchmark + Target Change	Comparison of actual spending to Target	Penalty for spending in excess of Target	100% of spending above Target	None

ACCOUNTABILITY FOR UTILIZATION/SPENDING IN CURRENT CMMI APMS										
APM	Measure(s)	Measure Type	Target Type	Benchmark	Target Change	Target	Performance Assessment	Type of Payment Adjustment	Amount of Penalty	Limit on Penalty
Oncology Care Model (Track 2)	Episode Spending: Total spending (except for a portion of Part D drug spending) during six months after initiation of chemotherapy, capped at 99th percentile	Population-Level Measure	Benchmark-Based Target	Combination Benchmark: (1) Counterfactual: Estimated spending for each cancer patient based on risk-adjusted national averages; (2) Prior Performance: Adjustment for oncology practice's past spending compared to other practices; (3) Comparison Group: Trend and use of novel therapies for all practices not in the APM	Desired Change: 2.75% reduction in spending compared to benchmark	Benchmark + Target Change	Comparison of actual spending to Target	Penalty for spending in excess of Target Bonus for spending below Target	100% of spending above Target	Penalty limited to 20% of Benchmark

**TABLE 13 (CONTINUED)
ACCOUNTABILITY FOR UTILIZATION/SPENDING IN CURRENT CMMI APMS**

APM	Measure(s)	Measure Type	Target Type	Benchmark	Target Change	Target	Performance Assessment	Type of Payment Adjustment	Amount of Penalty	Limit on Penalty
Comprehensive Primary Care Initiative Plus (CPC+)	Utilization: Risk adjusted actual number of hospital admissions compared to expected number of admissions					80th percentile of each measure calculated for eligible patients receiving care from non-CPC+ practices	Comparison of practice performance on each measure during year to Targets		% of Performance-Based Incentive Payment (either \$2.50 or \$4.00 per patient per month) based on actual performance on the measures compared to percentiles of the comparison group, with double-weighting for hospital admission measure	Limited to 50% of Performance-Based Incentive Payment paid to the practice
	Utilization: Risk-adjusted number of emergency department visits that do not result in a hospital admission compared to expected number of visits	Population-Level Measure	Comparison Group Target	N/A	N/A			Penalty for utilization higher (worse) than Target		

**TABLE 13 (CONTINUED)
ACCOUNTABILITY FOR UTILIZATION/SPENDING IN CURRENT CMMI APMS**

APM	Measure(s)	Measure Type	Target Type	Benchmark	Target Change	Target	Performance Assessment	Type of Payment Adjustment	Amount of Penalty	Limit on Penalty
Medicare Shared Savings Program (MSSP)	Total Cost of Care: Partially risk-adjusted total spending (except for Part D drugs) per assigned beneficiary during the calendar year	Population-Level Measure	Benchmark-Based Target	Prior Performance: Risk-adjusted spending in previous 3 years for beneficiaries who were assigned, increased by savings achieved previously	Comparison Group Change: Increase in national average beneficiary spending per beneficiary during initial performance period; increase in regional average beneficiary spending per beneficiary during later performance periods	Benchmark + Target Change	Track 1: Comparison of spending to 96.1%-98% of Target (% depends on # of patients assigned)	Bonus Only	None	N/A
							Other Tracks: Determination of whether actual spending is greater than 100.5%-103.9% of Target or less than 96.1%-99.5% of Target	Penalty for Increased Spending Bonus for Savings	Track 1+: 30% of amount of actual spending above Target Track 2: 40%-60% (depending on quality scores) of amount of actual spending above Target Track 3: 40%-75% (depending on quality scores) of amount of actual spending above Target	10% of Target or 8% of ACO Part A/B revenue, whichever is less 15% of Target 20% of Target

**TABLE 13 (CONTINUED)
ACCOUNTABILITY FOR UTILIZATION/SPENDING IN CURRENT CMMI APMs**

APM	Measure(s)	Measure Type	Target Type	Benchmark	Target Change	Target	Performance Assessment	Type of Payment Adjustment	Amount of Penalty	Limit on Penalty
Comprehensive Care for Joint Replacement (CJR)	<p>Episode Spending: Spending on hospital admission for hip or knee replacement + Total spending (except for Part D drugs and specified exclusions) during 90 days after discharge, capped at 99th percentile</p>	Population-Level Measure	Benchmark-Based Target	<p>Prior Performance: Years 1-4: combination of (1) prior performance of participant hospital and (2) average prior performance of all hospitals in multi-state region;</p> <p>Years 5+: Average prior performance of all hospitals in multi-state region</p>	<p>Desired Change: 3% reduction in spending if quality is acceptable or less than acceptable; 2% reduction in spending if quality score is good; 1.5% reduction in spending if quality score is excellent</p>	Benchmark + Target Change	<p>Comparison of actual spending to Target; aggregated across all eligible patients</p>	<p>Penalty for spending in excess of Target</p> <p>Bonus for spending below Target if quality is acceptable or better</p>	100% of spending above Target	Penalty limited to 20% of Target

**TABLE 14
ACCOUNTABILITY FOR QUALITY IN CMMI APMs**

APM	Measure(s)	Measure Level	Target Type	Target	Performance Assessment	Type of Payment Adjustment	Amount of Penalty	Limit on Penalty
BPCI-Advanced	All Episodes: Total readmission rate at hospital							
	All Episodes: % of physician's patients who have an advanced care plan							
	Inpatient Episodes: AHRQ Patient Safety Indicators	Population-Level	Competitive Target	Maximum Performance Among BPCI-Advanced Participants	% of Quality Target achieved for each quality measure, averaged over all measures	Penalty based on population-level performance against Quality Target	Proportional to bonus earned if total spending is below Spending Target. No penalty for poor quality if no net savings were generated	Maximum of 10% of amount by which actual spending is lower than Spending Target
	Some Procedures: Physician use of antibiotic							
	Hip/Knee Surgery: Complication rate							
	CABG: Mortality Rate							
	AMI: Excess Days in Acute Care							

**TABLE 14 (CONTINUED)
ACCOUNTABILITY FOR QUALITY IN CMMI APMS**

APM	Measure(s)	Measure Level	Target Type	Target	Performance Assessment	Type of Payment Adjustment	Amount of Penalty	Limit on Penalty
Oncology Care Model (Track2)	3 claims-based measures applicable to all patients	Population-Level	Comparison Group Target	80th percentile nationally	Percentile of actual measure	Penalty based on population-level performance compared to Quality Target	Performance on all measures are combined and compared to maximum possible score, then assigned one of 4 performance multipliers (0%, 50%, 75%, 100%). Performance multiplier applied to amount by which total spending is below Spending Target, if any, to determine final bonus for spending. No penalty for poor quality if no net savings were generated.	Maximum of 100% of any savings generated
	1 survey-based patient experience measure	Population-Level	Comparison Group Target	1 standard deviation above national mean	Actual score compared to national mean			
	4 registry-based measures applicable to all patients	Population-Level	Competitive Target	Undefined	Actual score compared to other practices			
	4 registry-based measures applicable to specific types of cancer	Population-Level	Competitive Target	Undefined	Actual score compared to other practices			

**TABLE 14 (CONTINUED)
ACCOUNTABILITY FOR QUALITY IN CMMI APMS**

APM	Measure(s)	Measure Level	Target Type	Target	Performance Assessment	Type of Payment Adjustment	Amount of Penalty	Limit on Penalty
Medicare Diabetes Prevention Program (MDPP)	Participant weight	Patient-Level	Status Quo + Goal-Based Change			First 6 Months: Bonus for Patient-Level Performance	\$160 bonus when Quality Target first achieved	
				5% reduction in weight from baseline	Current weight minus baseline weight	Second 6 Months: Penalty for Patient-Level Performance	\$45 reduction in payment if Quality Target has not been met or maintained	Fixed penalty
				9% reduction in weight from baseline	Current weight minus baseline weight	Second Year: Outcome-Based Payment	No payment unless Quality Target has been met or maintained	
						Bonus for Patient-Level Performance	\$25 bonus when Quality Target is first achieved	

TABLE 14 (CONTINUED) ACCOUNTABILITY FOR QUALITY IN CMMI APMS								
APM	Measure(s)	Measure Level	Target Type	Target	Performance Assessment	Type of Payment Adjustment	Amount of Penalty	Limit on Penalty
Comprehensive Primary Care Initiative Plus (CPC+)	CAHPS Experience of Care (average of 6 sub-measures)	Population-Level	Comparison Group Target	80th percentile of national performance on each measure for practices reporting through EHRs for MIPS	Comparison of practice performance during year to Quality Target	Penalty for Population-Level Performance worse than Quality Target	% of Performance-Based Incentive Payment (either \$2.50 or \$4.00 per patient per month) based on actual performance on the measures compared to percentiles of the comparison group.	Up to 100% of Performance-Based Incentive Payment paid to the practice
	9 electronically submitted PQRS/MIPS quality measures: 1. Blood pressure control 2. HbA1c control + 7 measures selected from a set of 17 measure options	Population-Level	Comparison Group Target	80th percentile of national CAHPS scores	Comparison of practice performance during year to Quality Target	Penalty for Population-Level Performance worse than Quality Target		

TABLE 14 (CONTINUED) ACCOUNTABILITY FOR QUALITY IN CMMI APMs								
APM	Measure(s)	Measure Level	Target Type	Target	Performance Assessment	Type of Payment Adjustment	Amount of Penalty	Limit on Penalty
Medicare Shared Savings Program (MSSP)	31 Quality Measures	Population-Level	Status Quo - Comparison Group	90th percentile of comparison group	Comparison of performance to Quality Target	Penalty for Population-Level Performance proportional to Penalty or Bonus for Spending	If spending is below Spending Target: % reduction in shared savings payment based on average % of Quality Target achieved, with adjustment for improvement If spending is equal to Spending Target: no penalty for poor quality If spending is greater than Spending Target in Track 2&3: % increase in penalty for Spending based on % of Quality Target achieved	Up to 100% of shared savings payment (if savings compared to Spending Target was achieved)
			Status Quo - Comparison Group	"Significant improvement"	Comparison of performance to Quality Target			

**TABLE 14 (CONTINUED)
ACCOUNTABILITY FOR QUALITY IN CMMI APMS**

APM	Measure(s)	Measure Level	Target Type	Target	Performance Assessment	Type of Payment Adjustment	Amount of Penalty	Limit on Penalty
Comprehensive Care for Joint Replacement (CJR)15	(1) Risk-standardized complication rate (2) HCAHPS patient experience measure	Population-Level	Status Quo - Comparison Group	90th percentile of all hospitals	Hospital performance compared to Quality Target and assigned to one of four categories: (1) Excellent (2) Good (3) Acceptable (4) Less Than Acceptable	Penalty for Population-Level Performance proportional to Spending Benchmark	If quality is less than acceptable and spending is below Spending Target, no bonus payment	Up to 20% of Spending Benchmark
			Status Quo - Comparison Group	Two decile improvement from prior performance relative to Comparison Group			If quality is at least acceptable, lower quality reduces Spending Target by 0.5%-1.5% of Spending Benchmark, which increases likelihood and size of penalty for high spending and reduces likelihood and size of bonus for low spending	Up to 1.5% of Spending Benchmark

Use of Component #3 in Medicare APMs

Table 14 shows that in most of the major Medicare APMs (the Medicare Shared Savings Program, the Oncology Care Model, Bundled Payments for Care Improvement – Advanced, and the Comprehensive Care for Joint Replacement program), there is no direct penalty if a provider delivers poor quality care to patients. The only “penalty” is that if the provider qualifies for a shared savings bonus or payment reconciliation amount because spending was held below the Spending Target, the bonus or reconciliation amount is reduced based on the quality score. However, if the provider does not reduce spending below the Target, there is no penalty for delivering low quality care.

In addition, these APMs only use population-level quality measures focused on a narrow range of quality issues, so it is possible that providers could receive bonuses for reducing spending even though some patients received poorer quality care.

In the Comprehensive Primary Care Plus (CPC+) APM, a primary care practice can be penalized for poor quality care – it has to return a portion of the Performance-Based Incentive Payment if performance on the quality measures is poor compared to other primary care practices. However, CPC+ only uses population-level quality measures, so it is possible that some patients could receive poorer quality care as long as there were offsetting improvements in care for other patients.

The only Medicare APM that is using patient-level quality measures and an Outcome-Based Payment is the Medicare Diabetes Prevention Program, where payments in the second year are only made if the participants maintain a 5% weight loss from their baseline weight when they first entered the program.

Use of Component #4 in Medicare APMs

Most of the Medicare APMs determine the eligibility of patients exclusively or primarily using retrospective attribution methodologies driven by fee-for-service utilization, rather than allowing providers and patients to determine whether the patient is eligible and wants to participate in the APM. As described in Section VI.D, even in the Oncology Care Model, where an oncology practice can bill CMS for a Monthly Enhanced Oncology Services (MEOS) payment when the practice begins delivering chemotherapy to a patient, CMS will recoup the payment if the CMS attribution methodology determines the patient did not receive the majority of cancer-related evaluation & management services from that practice.

A few Medicare APMs do allow providers and patients to determine eligibility and participation. In the Initiative to Reduce Avoidable Hospitalizations among Nursing Facility Residents – Payment Reform, patient eligibility is determined by a physician or clinician at the participating nursing facility who confirms that a patient has a qualifying diagnosis and will receive treatment for that diagnosis.²⁰⁵ The Next Generation ACO program uses a claims-based attribution methodology but it allows that to be overridden if a patient voluntarily “aligns” with the ACO.

b. Better Ways to Design APMs

Table 15 shows four general APM designs that are better than the APMs currently being used by CMS and other payers. They use the options described in Sections VI.A, VI.B, VI.C, and VI.D to create increasing levels of flexibility and accountability in the delivery of services. Although other combinations are possible, it is likely that one of these designs will be appropriate in most situations where an APM is needed. The four designs are:

- **Accountable Payment for Service.** A provider receives a new or revised payment for delivering a specific service to patients, and the payment is reduced if targets for spending on specific services and performance on quality measures are not achieved.
- **Accountable Bundled Payment.** A provider or team of providers receives a bundled payment to enable delivery of a group of services to patients or to treat a particular condition, and the payment is reduced if targets for spending on specific services and performance on quality measures are not achieved.
- **Outcome-Based Payment.** A provider is only paid for a service or group of services if standards or targets for quality and spending are achieved.
- **Bundled/Warranted Payment.** A provider or team of providers receives a bundled payment to deliver a group of services to patients, and the provider team is responsible for using the payment to cover the costs of necessary services and also to pay for avoidable services or services needed to address complications of treatment.

The Appendix describes several detailed examples of how these Alternative Payment Model designs could be used to address different types of opportunities for reducing spending and improving quality. These include:

- **Payment for a High Value-Service.** This example describes how an Accountable Payment for Service could support enhanced care management services for patients with chronic diseases.
- **Condition-Based Payment for an Acute Condition.** This example describes how a Bundled/Warranted Payment could support better care for pregnant women and their babies.
- **Condition-Based Payment for a Chronic Condition.** This example describes how a Bundled/Warranted Payment could replace current fee-for-service payments for managing a chronic condition.

These examples are not intended to represent the full range of APMs that are needed to address all opportunities for improvements in healthcare, nor are they necessarily the best approaches in every community. However, all of the examples represent better ways of addressing important opportunities than most of the APMs that have been developed to date.²⁰⁶

**TABLE 15
COMPONENTS OF WELL-DESIGNED APMs vs CURRENT APMs**

	Current APMs		Well-Designed APMs			
	Shared Savings	Population-Based Payment	Accountable Payment for Services	Accountable Bundled Payment	Outcome-Based Payment	Bundled/Warrantied Payment
Component #1: Adequate Payment for Needed Services	No change in FFS	Flexible payment for each patient; higher amounts for some but not all needs	Payments for new high-value service(s) and/or higher payments for existing service(s)	Bundled payment for group of services from a provider team	Payments for new high-value services and/or higher payments for existing services	Bundled payment for group of services from a provider team
Component #2: Accountability for Spending	Penalty for increase in total cost of care	Fixed payment regardless of services needed or delivered	Penalty if spending controllable by provider exceeds target		Penalty if spending controllable by provider exceeds target	
Component #3: Accountability for Quality	None	Penalties for poor performance on population-level quality measures	Penalty if quality controllable by provider falls short of target for individual patient	Penalty if quality controllable by provider falls short of target for individual patient	No payment if quality standards are not met	Compensation for problems caused by failure to deliver high-quality care
Component #4: Patient Eligibility Determination	Attributed based on service utilization	Attributed based on service utilization	Patient selects provider team	Patient selects provider team	Patient selects provider team	Patient selects provider team

2. Ensuring the APM Design Supports the Business Case

The justification for designing an APM is based on (a) the existence of a business case for reducing spending and/or improving quality through delivering healthcare services in a different way (as discussed in Section IV.C) and (b) the presence of barriers in the current payment system (as discussed in Section V) that prevent services from being delivered in that way. Consequently, once a preliminary APM design has been developed that includes each of the components discussed in Sections VI.A, VI.B, VI.C, and VI.D, analyses should be performed to ensure that the APM design would:

- Remove or adequately mitigate the barriers in the current payment system to enable the desired services to be delivered; and
- Pay amounts for services and achieve levels of savings and quality that create a desirable business case for both payers and providers to implement the APM. This means:
 - ◆ The payments are adequate to cover the costs the providers will incur in delivering services; and

- ◆ The savings generated are sufficient to offset any increase in payments compared to the current system.

In general, these analyses require constructing a simulation model that includes the following components:

- Estimates of the costs of delivering the desired services to different numbers and characteristics of patients;
- Estimates of the parameters of the payment model, including:
 - ◆ The dollar amount of each type of payment;
 - ◆ The Targets for both utilization/spending and quality/outcomes; and
 - ◆ The penalties and bonuses based on performance, if any.
- Formulas for calculating the costs of service delivery and the amounts paid for services under multiple scenarios that include differing assumptions about:
 - ◆ the number and types of patients participating;
 - ◆ the number and types of services those patients will need; and

- ◆ the number and types of services the patients will actually receive, the quality of those services, and the outcomes the patients will achieve.

The scenarios examined should reflect a full range of realistic possibilities, including scenarios that have the potential to be unfavorable for payers, unfavorable for providers, and unfavorable for patients.

If the simulations indicate there would not be a business case for implementing the APM in one or more scenarios, then modifications may be needed to correct that. This could include using a different option for payment from among those described in Section VI.A, different amounts of payments for services, different or additional measures of spending and/or quality, different methods of setting Targets for those measures, changes in the mechanism used for adjusting payments based on performance, or changes in the eligibility standards for patients.

Example: One study found that the amounts of payment under the Medicare Diabetes Prevention Program would not be sufficient to cover the costs of the program with the levels of success on outcomes that were achievable for the types of patients in the program.²⁰⁷

However, it is unlikely that any APM can guarantee that payments will be greater than costs, spending will be lower, and outcomes will be better in all possible scenarios. The same scenarios should be analyzed assuming that no changes are made in the payment system so that the APM can be compared to the status quo, rather than to a non-existent ideal. If the advantages of the scenarios in which the APM improves on the status quo are deemed to outweigh the disadvantages of the scenarios in which it does not, then it may make sense to implement the APM rather than retaining the status quo, even though the APM is not “perfect.”

In many cases, it may be difficult to accurately estimate the costs of delivering services in different ways because the services have never been delivered in those ways before, and it may be difficult to determine what parameters of the payment model are feasible until the services are actually delivered. In these situations, the scenarios should also include ranges of potential values for those costs and parameters, and simulations can be performed to determine whether the APM is likely to be successful despite this uncertainty or whether additional analysis or pilot testing is needed to make more accurate estimates. Section VIII.B will discuss how a beta-testing process for APMs could be created in order to refine and finalize parameters for an APM.

3. Creating Alternative Versions of the Alternative Payment Model

a. Evolving the APM Design Over Time

Since there are different options in each component that can be used to achieve similar goals, it is also possible to use one set of options for the initial APM design, and then transition the design to use a different set of options at a later point in time. This would likely be particularly desirable for very innovative APMs, such as those that are designed to support major changes in the way a service is delivered, those designed to support simultaneous changes in multiple types of services, and those designed to improve patient outcomes rather than narrowly defined process measures. In these cases, it will be more difficult to define the correct payment amounts for services and to set good Targets for spending and/or quality, and as a result, both providers and patients will likely be more concerned if the APM includes significant penalties for failure to achieve the Targets.

For example, the Accountability Components could evolve through two or more of the following stages:

1. Penalties/Bonuses for Population-Level Targets on Spending and/or Quality Measures
2. Penalties/Bonuses for Patient-Level Targets on Spending and/or Quality Measures
3. Outcome-Based Payments
4. Bundled/Warranted Payments

The use of Patient-Level Targets is more desirable from the patient’s perspective but they are more challenging for a provider to achieve than a Population-Level Target, so providers could begin with a Population-Level Target while care delivery methods are being refined, and then transition to Patient-Level Targets. An Outcome-Based Payment involves the full loss of payment when a Target is not achieved, so a provider could begin with smaller penalties under the Penalty option before moving to an Outcome-Based Payment. Finally, since the Bundled/Warranted Payment can result in a penalty greater than the provider’s payment, beginning with an Outcome-Based Payment would create less financial risk for the provider.

The payment changes made to remove the barriers to delivering services could evolve through two or more of the following stages:

1. Fees for Currently Unpaid Services or Increased Fees for Underpaid Services
2. Bundled Payments for Groups of Services
3. Condition-Based Payments
4. Multi-Provider Bundled/Condition-Based Payments

The larger the payment bundle, the greater the difficulty in determining the “right” payment amount. To address this, a provider could begin with unbundled payments for individual services in order to better understand the costs and benefits of delivering new and modified services. This would enable better decisions to be made about what should be included in a bundle and what payment amount would be appropriate.

b. Creating Different Designs for Different Providers and Communities

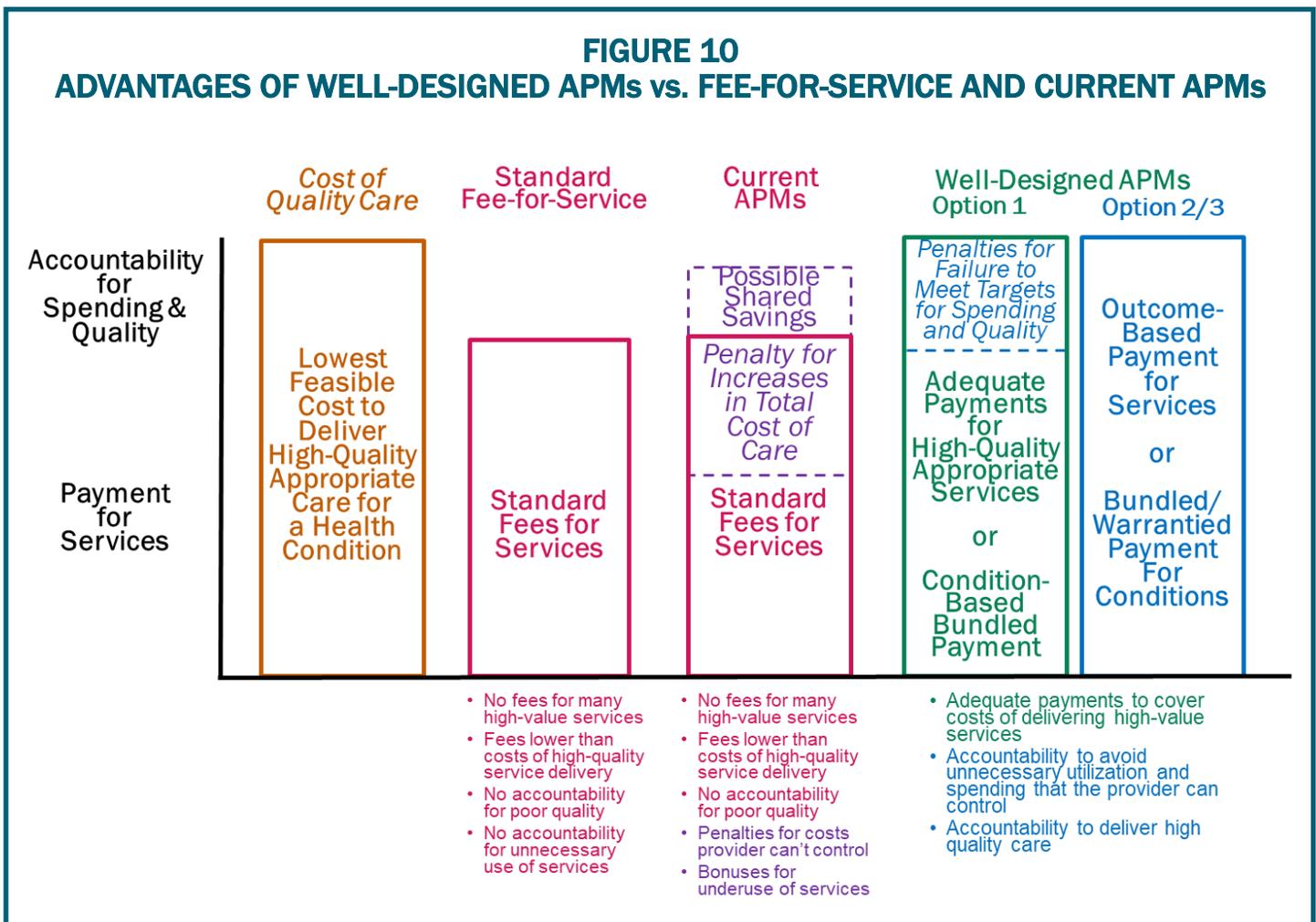
The ability to use different options in each component to achieve similar goals also provides the opportunity to use different versions of the APM to better match the resources and capabilities of different providers in different communities. For example:

- Providers that have a small number of patients will inherently be at greater risk of random variation on both spending and outcomes, so different methods of making performance-based adjustments to payments could be used for smaller providers vs. larger providers. For example:
 - ◆ Option 4 discussed in Sections VI.B.4 and VI.C.4 could be used for very small providers or for providers with a very small number of eligible patients. The same performance measures and Targets could be used as for larger providers, but a customized method of evaluation would be created to determine whether the small provider failed to achieve a Target because of random or unavoidable reasons.
 - ◆ Option 1 (penalties) could be used instead of Option 2 (outcome-based payment) for smaller providers, since Option 1 provides more flexibility to limit the magnitude of penalties.

A team of providers that can deliver all of the planned services needed for care under the APM will be in a better position to implement a Bundled/Warranted Payment, whereas providers who order some of the services from providers located in different communities and rely on them to help avoid unnecessary services and complications will likely be more comfortable implementing an Outcome-Based Payment or a Penalty approach.

Example: Large academic medical centers often perform procedures on patients who live many miles away and travel to the hospital for the procedure. If the patients need post-acute care, they will likely receive it in their home community rather than in the city where the medical center is located. However, this makes it difficult for the medical center to manage a bundled payment for the full episode of care, because there will be a large number of post-acute care providers from a broad geographic area involved, even if the episodes all begin in the same hospital.

FIGURE 10
ADVANTAGES OF WELL-DESIGNED APMs vs. FEE-FOR-SERVICE AND CURRENT APMs



VII. OPERATIONALIZING THE ALTERNATIVE PAYMENT MODEL

STEP 5 Determine how payers & providers can operationalize the APM

A. Change Payments for Services

- Create CPT/HCPCS codes or modifiers
- Define correct coding rules
- Define time periods for service bundles
- Define default allocations of payments in bundles

B. Determine Eligibility of Patients

- Use CPT/HCPCS codes to indicate patient eligibility

C. Measure Performance on Spending & Quality

- Create ICD and CPT codes to identify related services
- Create HCPCS codes to record quality performance

D. Adjust Payments for Performance

- Create HCPCS codes to record achievement of Targets

Once decisions have been made about the options for each of the components of the Alternative Payment Model, additional details are needed in order to operationalize the APM. Mechanisms are needed for making determinations as to whether and how much the providers participating in the APM should be paid for specific patients in specific situations, and these mechanisms need to be feasible for payers and providers to implement.

Many payers have cited operational challenges as an excuse for refusing to implement desirable APMs. Instead, they have pursued APMs that are easy for them to operationalize but problematic for providers or patients. For example, most of the APMs that have been implemented to date use “shared savings” approaches, not because that is the ideal way to structure a payment model, but because it is extremely easy for payers to implement a payment model that requires no changes in the way payments are made for services. However, this simple-to-implement approach is likely to be ineffective in significantly reducing spending or improving quality because it fails to directly address the many barriers in the fee-for-service system that prevent delivering care in better ways.

At least in the near term, any APM will need to operate in parallel with the existing payment system rather than replace it. Consequently, the APM will be easiest to operationalize if it can use existing billing systems, claims payment systems, and data collection mechanisms to the maximum extent possible. Fortunately, even though current claims forms and coding systems were designed for the current fee-for-service system, the same forms and systems can also be used to operationalize most aspects of APMs. The key is to translate the structure of the APM into the “language” of claims payment systems – procedure codes, modifiers, diagnosis codes, edit processes, etc.

A. Changing Payments for Services

Section VI.A describes multiple options for paying providers differently than they are paid in current fee-for-service systems. Different methods are needed to operationalize each of these ways of paying for services within existing billing and claims payment systems.

1. Payment for a Currently Unpaid Service

If the APM will pay a provider a fee for delivery of a service for which there is no current payment (Option 1 in Section VI.B), then a new billing code will need to be created so the provider can bill a patient or payer for the service. In most cases²⁰⁸, this will be a Current Procedural Terminology (CPT®)²⁰⁹ or Healthcare Common Procedure Coding System (HCPCS)²¹⁰ code.

In some cases, there may already be a code that describes a particular service, but there is no payment associated with that code. All that is needed in these cases is for the payer to agree to pay a provider that is participating in the APM for that code when it is submitted on a claim. For example, there are CPT codes for telephone calls between physicians and patients, but they are not currently eligible for payment under Medicare, so if an APM is intended to provide payments to physicians for the types of phone calls described in the CPT codes, the physician could bill a payer or patient using the CPT codes.²¹¹

If there is not an existing CPT or HCPCS code that accurately describes the service that will be paid under the APM, a new code will need to be created. CPT Codes are created by the American Medical Association’s CPT Editorial Panel, and there is an established process for requesting the creation of new CPT codes.²¹² HCPCS Level II codes are established by the Centers for Medicare and Medicaid Services (CMS) through the CMS HCPCS Workgroup.²¹³ “Permanent” HCPCS Level II

**TABLE 16
OPERATIONALIZING COMPONENT #1**

Payment Option	Method of Operationalizing the APM
<i>Paying for an Unpaid Service</i>	
1. Pay a fee for the service	Create a CPT®/HCPCS code for the service
2. Bundled payment for a group of services	Increase the payment for an existing CPT/HCPCS code in the bundle; Use a modifier with an existing CPT/HCPCS code to trigger bundled payment; Create a new bundled CPT/HCPCS code to replace existing codes
<i>Aligning Payment with the Cost of a Service</i>	
3. Higher payment for the service	Pay more when the service is delivered by APM participants; Use a CPT/HCPCS modifier when the service is delivered to an APM patient
4. Payment stratified by phase of care	Create a family of CPT/HCPCS codes or a series of modifiers that distinguish the phase of care in which the service is delivered
5. Payment stratified by patient characteristics	Create a family of CPT/HCPCS codes or a series of modifiers that distinguish the characteristics of the patient to whom the service was delivered
6. Condition-based payment	Create a CPT/HCPCS code that describes the patient’s condition instead of the service that was delivered
7. Standby capacity payment	Make a payment to the standby provider for each insured member who lives in a geographic area; Make a payment to the standby provider for each patient for whom any provider is receiving a condition-based payment or a payment for an alternative service
8. Volume-based payment adjustment	Use a modifier or add-on CPT/HCPCS code to indicate that a service was delivered by a low-volume provider and requires higher payment
9. Outlier payment	Use a modifier or add-on CPT/HCPCS code to indicate that a service was delivered to a patient who required additional time or cost
10. Cost-based payment	Obtain documentation of a provider’s cost for delivering the service and allocate the costs among patients and payers
11. Multi-component payment	Combine two or more of the methods described above
<i>Enabling Control/Coordination of Other Providers’ Services</i>	
12. Multi-provider bundled payment	Make the payment to an organization controlled by participating providers; Pay individual providers based on standard FFS amounts and reconcile the difference through an organization controlled by one or more providers; Use a CPT/HCPCS modifier to indicate the provider should be paid a proportion of the bundled payment rather than a fixed amount
<i>Modifying Patient Cost-Sharing</i>	
13. Modified first dollar cost-sharing	Change cost-sharing requirements for specific CPT/HCPCS codes; Use a modifier to indicate that cost-sharing is different for an APM patient
14. Last-dollar cost-sharing	Require the patient to pay the difference between the provider’s charge and the payer’s maximum payment amount for the service

codes are defined in order to enable payment for drugs and medical devices, since CPT codes describe only services, not products. However, there are also many HCPCS Level II “temporary” codes that describe specific services delivered by physicians or other healthcare practitioners for which there is not a CPT code.²¹⁴

Even though there are over 9,000 CPT codes, CMS has created hundreds of temporary “G-codes” under HCPCS to describe services for Medicare beneficiaries for which there is not a CPT code, private payers have created hundreds more “S-codes” under HCPCS for services delivered to commercially-insured patients for which there is no CPT code, and state Medicaid agencies have created dozens of HCPCS “T-codes” for services delivered to Medicaid beneficiaries for which there is no CPT code.²¹⁵

CMS has created a number of G-codes specifically to enable billing and payment for services as part of Alternative Payment Models, and private health plans have used S-codes to pay for services under alternative payment models they have created.²¹⁶ For example,

- CMS created 9 new G-Codes (G9481-G9489) to enable physicians and clinicians to be paid for making in-home “visits” to new and established patients using telehealth technologies if the patients are part of a Next Generation ACO.
- CMS created the HCPCS code G9678 to enable oncologists participating in the Oncology Care Model to bill for the “monthly enhanced oncology services (MEOS) payment” that they can receive on a monthly basis for each eligible patient.

The existence of a code does not necessarily mean that the service will be paid for by Medicare, Medicaid, or a commercial health insurance plan for all providers or patients. It merely provides a mechanism by which a physician, hospital, or other healthcare provider can bill for the service if the payer has agreed to pay for that service when it is delivered to an eligible patient by an eligible provider.

When a provider submits a claim with a code that is being paid for only under an APM, the provider would be certifying that it delivered the service that is described by the code. If the APM specifies that the service can only be delivered to certain types of patients (e.g., patients with a particular disease), then the provider would have the responsibility of verifying and documenting that a patient met those eligibility criteria before submitting the code on a claim form requesting payment. This is no different than what exists under standard fee-for-service billing – when a provider submits a CPT code on a claims form, the provider is certifying that the service described by the CPT code was delivered to a patient for whom the service was appropriate.

The payer would still need to verify that the provider is participating in the APM and that the provider is eligible to be paid for delivering that specific service to that specific patient under the APM. This is also no different than what exists under the current fee-for-service system, where the payer would verify that the provider is participating in the payer’s insurance program, and that the patient is covered for the specific service that the provider delivered.

New CPT and HCPCS codes are created every year, so every payer has to have the capability of paying for new codes, and any provider’s billing system should have the ability to bill for new codes. Consequently, there should not be any significant operational barriers preventing either providers or payers from participating in an APM that requires payment for a new billing code.

It is important to recognize that even though a provider bills and is paid for a new service under the APM with the same types of service/billing codes used in standard fee-for-service payment, paying for a code under the APM is different from simply adding that code to the standard list of fee-for-service codes. Under the APM, the provider delivering the service would be accountable for spending and quality under whatever mechanism is defined in the APM, whereas under standard fee-for-service payment, there would be no similar accountability.

2. Different Payment Amount for an Existing Service as Part of the APM

If the APM needs to pay a higher amount for a service than would ordinarily be paid under the standard fee-for-service payment system (Option 3 in Section VI.A), this could be operationalized in several ways:

- If an individual provider that participates in the APM will *only* deliver the service as part of the APM, the provider could use the same billing code as other providers do, but the payer would simply pay a different amount to the provider based on their participation in the APM. Medicare and other payers currently pay different amounts for the same service to different providers or different types of providers, so it should be feasible to create a different payment amount for those providers participating in the APM.
- If the provider will deliver the service to some patients as part of the APM but the provider will also deliver the same service to other patients who are not part of the APM, there needs to be a way to notify the payer when the service is being delivered as part of the APM and when it is not. The CPT/HCPCS coding system has a standard mechanism for indicating that the same service was delivered under different circumstances that justify different payment amounts. This is done by appending a two-digit alphanumeric modifier to the five-digit service code. There are many such modifiers already.

Example: Modifier 22 is appended to a CPT procedure code to indicate that “the work required to provide a service is substantially greater than typically required,” and Modifier 63 is appended to indicate that the procedure was performed on an infant weighing less than 4kg.

- In some cases, an APM might be designed to pay for a new service only when it is delivered along with another service, but it would not always be delivered with the other services. For example, translation services may be needed for patients who do not speak English or who have other language difficulties in order to enable an office visit with a physician or other clinician to be successful. In these cases, an “add-on

code” could be created describing a service that is only paid for if another service is also delivered (e.g., language translation services delivered in conjunction with an office visit). There are already many types of add-on codes in the CPT/HCPCS code list that are used to modify payments for services under Medicare and other types of insurance. CMS regularly updates the list of CPT/HCPCS codes which must also be billed in order for payment to be made for the add-on code.²¹⁷

Example: CPT Code 59525 is used to indicate that a hysterectomy was performed after a Cesarean delivery and it must be accompanied by a CPT Code indicating that a C-Section was performed (59510, 59514, 59515, 59618, 59620, or 59622). If the hysterectomy is performed outside of a childbirth situation, CPT Codes 58150-58294 are used.

3. Payments Stratified by Phase of Care

If an APM needs to pay different amounts for what is ostensibly the same service when it is delivered in different phases of care (i.e., Option 4 in Section VI.A), this could be operationalized by defining two or more codes for the service, with each code assigned to a different phase of care.

Example: In the Medicare Physician Fee Schedule, there is a higher amount of payment for the same Evaluation and Management service when the physician sees a new patient rather than a patient they have seen previously (an “established” patient).

Example: In the Medicare Diabetes Prevention Program, patients attend a series of one-hour “structured health behavior change sessions” designed to promote weight loss through healthy eating and physical activity. The amount of payments for these sessions differ depending on how long the patient has been participating in the program as well as whether the patient has lost weight and how many total sessions they have attended. During the first six months of participation, providers bill using HCPCS codes G9873, G9874, and G9875 to receive a payment of \$25 for the first session, \$50 for the next three sessions, and \$40 for the next five sessions (\$8 per session). During months 7-12, providers bill using HCPCS codes G9876, G9877, G9878, and G9879 to receive a \$15 payment for two sessions (\$7.50 per session), unless the patient achieves a 5% weight loss, in which case the payment increases to \$60 (\$30 per session). During months 13-24, providers bill HCPCS codes G9882, G9883, G9884, and G9885 for each pair of sessions the patient attends if the client has achieved and maintained a 5% weight loss.²¹⁸

Many current APMs and proposed APMs have failed to recognize the differences in costs for different phases of care and have either attempted to pay for all phases in the same way or to pay differently for only one phase of care and retain the current payment system for others.²¹⁹ These APMs cannot achieve one of the most im-

portant goals of an APM – ensuring that payments support the most appropriate services that patients need.

If multiple APMs are stratifying payments by phase of care and the same definitions of care phases are being used, an alternative would be to define a standard set of modifiers to indicate the phase of care in which the service was delivered, and the modifier would then be used to signal the need for a different payment amount.

4. Payments Stratified by Patient Characteristics

In many APMs, there will be a need to pay different amounts for a service when it is delivered to patients with different characteristics. Payments for a service that are stratified into two or more discrete categories of patients (Option 5 in Section VI) can be implemented in a claims payment system by defining a *family* of codes, one for each category of patients. Each category would be defined using one or more characteristics of patients that affect the cost of delivering the service, and then a separate CPT/HCPCS code would be assigned to each category. There would not need to be a separate code for each unique combination of patient characteristics; if the cost of delivering services was similar for patients with two different sets of characteristics, both combinations of characteristics could be assigned to the same category and the same code. A provider delivering the service would determine which category the patient qualifies for by comparing the patient’s characteristics to the criteria for each category, and then the provider would submit a claim for payment using the code that applies to that service and that category of patients.

This approach is already used for some types of services; for example:

- Nephrologists receive monthly payments to deliver care to patients who have end-stage renal disease (ESRD) and are receiving hemodialysis. There are four separate categories of CPT codes and different payments based on the age of the patient (<2 years of age, 2-11 years, 12-19 years, and 20 years of age or older).
- Medicare pays for hospital procedures using the Inpatient Prospective Payment System. For most procedures, there are 2-3 different Diagnosis Related Group (DRG) codes, with each of the codes representing patients with differing characteristics that affect the costs of treatment.²²⁰

As explained in Section VI, stratification of patients into discrete categories has a number of advantages compared to the use of “risk scores” calculated based on each individual patient’s characteristics. In addition, it is more difficult to operationalize an APM that uses risk scores because standard billing and claims payment systems are not designed to pay unique amounts for each individual patient based on a continuous variable such as a risk score. If it is necessary to use a risk scoring system, discrete categories based on ranges of risk scores can be defined.²²¹

Example: In the Medicare Comprehensive Primary Care Plus initiative, primary care practices receive a monthly Care Management Fee (CMF) for each eligi-

ble patient. The patient is assigned a risk score using the CMS Hierarchical Condition Categories (HCC) risk adjustment system, and that risk score is compared to the distribution of risk scores for all Medicare FFS beneficiaries in the same region who meet CPC+ eligibility requirements. The patient is then assigned to one of 4 or 5 tiers based on the percentile in which their risk score falls, and a higher CMF payment is paid for patients in tiers associated with higher risk scores. For example, for a patient whose risk score is below the 25th percentile, a Track 1 practice receives \$6 per month, whereas for a patient whose risk score is in the 75th percentile, the Track 1 practice receives \$30 per month. (Practices in Track 2 receive higher CMF payments in each track.)²²²

5. Payment for a New Service In Lieu of an Existing Service

In some cases, the APM will be paying for a new service with the expectation that it will be used *instead of* another service that the provider might have delivered, and not in *addition to* that other service. In these situations, a code would be needed for the new service, but the payer would need to ensure that if claims are submitted for the same patient with codes for both services, one or the other of the claims would be rejected.

This situation is not unique to an APM. The same issue arises in the current fee-for-service system when two different codes describe two different approaches to delivering a service or treating a condition, but both approaches cannot be used for the same patient. In these cases, it would be inappropriate for a provider to be paid for both codes. One of the goals of the “National Correct Coding Initiative (NCCI)” is to avoid paying for two service codes when only one is appropriate. Part of the NCCI is the Procedure-to-Procedure (PTP) Column One/Column Two Correct Coding edit file that is developed and maintained by CMS.²²³ CMS and other payers operate “claims editing” processes to identify codes on claims that violate these rules and then they deny payment for these codes. The NCCI rules would simply need to be expanded to define the types of codes that will not be paid for in addition to the APM service code.²²⁴

If there are some circumstances in which a payment should be made for the new service in addition to an existing service and other circumstances in which it should not, then a modifier could be created to distinguish those situations. In addition to signaling the need for different amounts of payment, modifiers are also used in the current FFS system to signal that a service should be paid for separately.

Example: *In the Medicare Physician Fee Schedule, Modifier 25 is used to indicate that a patient who was receiving a procedure or a visit with a physician also had a separate visit with a physician for a different purpose, so that an additional payment should be made for that visit.*

6. Bundled Payments for Planned Services

Operationalizing a bundled payment for planned services delivered by a single provider (Option 2 in Section VI.A) does not require radical changes in current billing and claims payment systems. In fact, there are a number of CPT/HCPCS codes that are explicitly defined as bundles of two or more other services.

Example: *CPT code 59400 describes a full bundle of routine obstetric care services, including prenatal care, vaginal delivery, and post-partum care. There are also separate codes describing each of the separate components of the bundle (59425 and 59426 for prenatal care, 59409 for the vaginal delivery alone, and 59430 for post-partum care). Ordinarily, an obstetrician will bill for the full bundle of services, but if different physicians provide the individual components, the separate codes enable them to be paid for their individual services.*

In addition to creating a billing code for the bundled payment, two things are needed:

- a. a definition of either (i) the types of services that are supposed to be delivered as part of the bundle or (ii) the goal that the bundled services are supposed to achieve;
- b. a set of rules as to whether and when codes for individual services can be used instead of or in addition to the bundled code. The billing and claims payment systems can implement these rules using the correct coding process described above for avoiding paying for both of two services that are supposed to be alternatives.

7. Payment for a Bundle of Services Over a Period of Time

If the services covered by the bundled payment will not all occur at the same time or on the same day, it will be necessary to specify the period of time in which the services are expected to occur. During this period of time, if a provider delivers a service that is included in the bundle, it would be presumed that no additional payment would be made for that service beyond the bundled payment unless the provider uses a modifier to specifically indicate that the service was unrelated to the bundled payment. After this time period, no modifier would be needed in order for a provider to be paid for the individual service, and another bundled payment could be billed and paid if another bundle of services is delivered.

There is already a mechanism for doing this in the physician payment system used by Medicare and most payers. Each CPT/HCPCS code can have an associated “global period” that defines the period of time in which the provider billing the code is expected to deliver the services associated with the bundle and not to bill separately for those services in addition to the bundled payment.

Example: Surgical procedures generally are assigned to one of three different "global periods:"

- *A 0-day global period, which is used for endoscopies and some minor procedures, where it is assumed that everything related to the procedure will be completed on the same day.*
- *A 10-day global period, which is used for most minor procedures. The time period is actually a total of 11 days – the day of the procedure and the ten subsequent days.*
- *A 90-day global period, which is used for major procedures. The time period is actually 92 days, including the day before the surgery, the day of the surgery, and the 90 days after the surgery.*

The physician (the surgeon or proceduralist) receives no separate payment for visits with the patient that are related to the procedure if the visits occur during the global period. If the visit with the patient is for an unrelated reason, a modifier is used with the billing code to indicate that it should be paid for in addition to the payment for the surgical procedure. However, if the patient has post-operative complications that require a return to the operating room, that is paid for separately. (The payment bundle only applies to the surgeon; other physicians are still generally paid separately for their services.)

The shorter the global period, the greater the risk that a provider could delay delivering an important service until after the global period ends so that it can be billed separately to generate additional revenue. However, the longer the global period, the more likely it is that a patient could have a new problem that is unrelated to the original bundle and requires delivery of one of the same services that is included in the bundle. In these cases, it may be better to pay for the individual services as they are delivered and then reconcile them against a bundled payment amount. This is described in more detail in subsection 8d below.

8. Bundled Payment for Services Delivered by Different Providers

Implementing a bundled payment is more complex if the services included in the bundle will be or can be delivered by two or more different providers, e.g., two different physicians, a physician and a hospital, a hospital and a skilled nursing facility, etc. (Option 12 in Section VI.A). In addition to a definition of what services are included and excluded from the bundle (similar to what is discussed in #6 above), a mechanism is needed to enable each of the providers to be paid for the portion of the bundle of services that they delivered. There are at least four different mechanisms that can be used to do this:

- a. paying the full bundled amount to one entity;
- b. making pre-defined allocations of a payment for subsets of services;
- c. making pre-defined allocations of a payment to providers based on their roles; or
- d. paying for individual services and retrospectively reconciling those payments against a budget

8a: Payment of the full bundled amount to one entity

Under this mechanism, one provider, or an entity representing multiple providers, bills for the bundled payment using the appropriate billing code. That provider/entity is paid the full bundled payment amount and it takes responsibility for compensating the other providers for the services they delivered. This mechanism requires that:

- The other providers would not bill the payer for their services unless the services are delivered for reasons outside of the scope of the bundled payment. The payer would need to have mechanisms to avoid paying for services included in the bundle if a provider accidentally submits a bill for those services, and a modifier would need to be created if providers will not always deliver a particular service through the bundle.
- The provider/entity that is receiving the payment has mechanisms for determining what the other providers did, deciding how much to pay them for it, and transmitting payment to them. This may require the creation of new service codes and definitions if the bundled payment will be used to support different approaches to service than what is defined in current CPT/HCPCS codes.²²⁵

If it is not clear in advance which other providers will be delivering services under the bundle and/or if a particular type of service could be delivered to the patient and be paid for separately from the bundled payment, a modifier will be needed to indicate that a service is to be paid for separately from the bundle even though the default would be for that service to be included in the bundle.

8b: Pre-defined allocations of payment for subsets of services

Under this mechanism, specific percentages of the bundled payment are assigned to two or more subsets of services. If one provider delivers all of the services, that provider receives the full bundled payment, but if two providers deliver different subsets of the services, each provider receives the pre-defined percentage of the total payment assigned to the subset of services they delivered.

Medicare and other payers already have a mechanism for doing this for global surgical fees, which are intended to represent a bundle of services – not just the surgery itself, but pre-operative care and up to 90 days of post-operative services. If the surgeon delivers all of those services herself, she simply bills for the appropriate global fee. However, in some cases, the surgeon will not be able to provide all of the post-operative care directly (e.g., if a patient receives surgery at a hospital located in a different community) and she may transfer part of the post-operative care to a different physician. In that case, the surgeon would bill for the surgical code with the "54" modifier code appended (indicating that she performed the surgery but not all of the other services), and the second physician would bill for the same surgical code with the "55" modifier ap-

pended (indicating that he provided postoperative management only), and each physician would indicate the proportion of the 90 day global period during which they were responsible for the patient's care. The Medicare Physician Fee Schedule has standard percentages for allocating the total bundled payment for the surgery between the pre-operative services, the surgery, and the post-operative services, so the post-operative percentage would be applied to the total payment to determine the total amount the surgeon and the other physician would be paid, and that amount would then be split between the surgeon and the second physician based on the proportions of the post-operative period they were each responsible for.

If the subsets of services that can be paid separately are defined adequately by existing service codes, then a modifier could be created to indicate that these services are being delivered as part of the overall bundle of services. Otherwise, one or more new service codes would need to be defined for the specific subsets of services so that providers could bill and be paid when they deliver one of those subsets.

8c: Pre-defined allocations of payment based on provider roles

Under this mechanism, specific percentages of the bundled payment would be assigned to providers based on the role they played in managing or delivering care under the bundled payment. Each of the providers could submit a bill using the same code describing the bundled service, but each would append a modifier to the code to indicate their role in delivering services as part of the bundle. The payer would then divide the overall bundled payment into shares for each provider based on a pre-defined proportion assigned to each modifier.

Under MACRA, Congress required the creation of "patient relationship codes" that physicians and other clinicians could use to define the nature of the role they were playing when they delivered a particular service. CMS has implemented this requirement by creating a set of modifiers that can be appended to a CPT/HCPCS code in order to indicate the physician's role.²²⁶ These modifiers are:

- X1: Continuous/Broad Services
- X2: Continuous/Focused Services
- X3: Episodic/Broad Services
- X4: Episodic/Focused Services
- X5: Only as Ordered by Another Clinician

If these modifiers can appropriately differentiate the different roles in a bundled payment, they could be used as described above, otherwise it may be necessary to develop alternative or additional modifiers to use in the APM.

8d: Retrospective reconciliation of billings for individual services

Under this mechanism, if a provider delivers a service that is included in the bundled payment, the provider bills the appropriate code for that service and they are

paid the standard amount for the service by the payer. At the end of whatever period of time is defined for the bundled payment, the payer sums all of the individual payments and compares the sum to the bundled payment amount. If the bundled payment amount is more than the sum of the individual payments that have already been made, an additional payment is made to one of the providers (or to an entity representing all of the providers); if the bundled payment is less than the sum of the individual payments, then one of the providers (or an entity representing all of the providers) is responsible for repaying the difference to the payer.

This mechanism is necessary in circumstances in which multiple providers may be billing for services included in the bundle but where (a) those providers are not all known in advance, or (b) some of the providers are not willing to have their payments come through other providers or a joint entity. In particular, if the patient's health insurance plan is unwilling or unable to require the patient to only obtain services from the providers who are participating in the APM, the payer will need to pay those other providers directly, and then deduct those payments from the bundled payment that would be paid to the APM participant entity.

There are three ways in which retrospective reconciliation can be operationalized:

- Withholds.** When the APM participants bill for the bundled payment, the payer initially pays the participants less than 100% of the bundled payment amount, i.e., the payer withholds a portion of the payment. If the payer receives claims for payment from other providers, it pays those claims and if the total of those payments is less than the withheld amount, the payer pays the difference to the APM participant.
- Delay in Payment.** The payer does not pay the APM participants anything until after a sufficient period of time has elapsed to be sure that all claims from other providers have been filed.
- Recoupment.** The payer pays the full amount of the bundled payment when the APM participants bill for it, but the APM participants must then repay the payer for any payments it made to other providers.

Example: In the Bundled Payments for Care Improvement Advanced (BPCI Advanced) APM, all providers – hospitals, physicians, skilled nursing facilities, home health agencies, etc. bill for their services as they are delivered and they are paid by Medicare using the standard methodology and payment amounts used for those services. CMS then adds up the total payments it has made for services associated with a particular Clinical Episode and compares that to the Target Price for the episode. If the total payments are less than the Target Price, the entity that is managing the episode receives an additional payment (called the Net Payment Reconciliation Amount, or NPRA) from CMS that is equal to the difference, and if the total payments are more than the Target Price, the entity must repay the difference to CMS.²²⁷

Prospective Payment vs. Retrospective Reconciliation

In order to distinguish it from retrospective reconciliation, the first mechanism (8a) is often referred to as “prospective payment,” even though the payment is made after the services are delivered. Under the first mechanism, the determination of how much each provider would receive is made after the bundled payment amount is paid, whereas under the fourth mechanism (8d), the individual providers are paid first and then one or more of the providers is responsible for reconciling any difference from the bundled payment amount. “Prospective payment” offers the greatest flexibility to deliver new types of services and to pay different amounts for existing services, since no provider’s payment is tied to any specific definitions of services or payment amounts. However, it requires either that all participating providers agree to be paid by another provider, rather than the patient’s health insurance plan, or that a new entity be formed to receive the payment. It also requires the payer to hold a different entity accountable for using payments appropriately.

9. Bundled Payment for Planned and Unplanned Services

The most challenging version of a bundled payment to operationalize is one that includes unplanned services as well as planned services. This is more difficult than a bundle based only on planned services because the provider who is delivering an unplanned service may not know that the patient has received planned services as part of a bundled payment, and so the provider’s bill for the service would not specifically indicate that it should be included in the bundled payment.

For example, if a bundled payment for a hospital procedure is designed to cover the cost of hospital readmissions that occur to treat complications of the procedure, a readmission could occur at a completely different hospital than the one where the procedure was performed. The physicians and hospital delivering care to the patient during the second admission will bill the payer for their services, but they may not realize or document that the admission was a “readmission” related to a previous procedure.

Consequently, a mechanism will need to be created for determining when an unplanned service would be considered as part of the bundle and how the payment for that service will be addressed by the bundle. This will generally require:

- creating a list of the types of unplanned services and associated diagnosis codes that the bundled payment would be expected to cover;
- defining the time period in which the unplanned services would be considered as part of the bundled payment;
- defining the mechanism for payment or reconciliation if one of the unplanned services is delivered by a provider other than the provider receiving the bundled payment.

Unless there is a reason to believe that the unplanned service will only be delivered by a provider that is delivering planned services as part of the bundle, payment for the unplanned service will likely need to be addressed through the retrospective reconciliation process described in subsection 8d.

10. Condition-Based Payments

In each of the situations described above, the payment is tied to the delivery of a specific service or combination of services that can be associated with a CPT code or HCPCS code describing the service(s). In contrast, under a condition-based payment (Option 6 in Section VI), the provider would be paid based on the type(s) of health condition(s) the patient has, rather than the specific services that are delivered. This provides greater flexibility to a provider or group of providers regarding the services they can deliver, but it makes it impossible to use a service-based code to trigger the payment.

Although ICD-10 diagnosis codes describing a patient’s conditions are currently recorded on claims forms, these codes do not provide an effective mechanism for billing a condition-based payment for several reasons:

- ICD-10 codes may not capture all aspects of a patient’s condition that are used to determine when the condition-based payment should apply.
- If the condition-based payment is designed to support management of a combination of problems rather than one specific disease, there is no mechanism for indicating on the claims form that (a) the listed ICD-10 codes collectively meet the eligibility requirement for the condition-based payment, and (b) the provider wishes to be paid for managing the patient’s care using the condition-based payment.
- Typical claims payment systems are not designed to determine the amount of payment based on an ICD-10 code or combination of codes without an accompanying CPT/HCPCS code. A claims payment system may require that an ICD-10 code be present in addition to a CPT/HCPCS code in order to document medical necessity for the service, but the amount of payment is based on the CPT code, not the ICD-10 code.

Consequently, in order to operationalize a condition-based payment in current billing and claims payment systems, a CPT/HCPCS code would need to be created for “treatment/care of a patient with condition X,” and the coding manual would need to include a definition of:

- the health condition or combination of conditions that the patient must have in order for the payment to be made;
- the minimum standard, if any, for the outcome that is to be achieved or the nature of the care or treatment that is to be delivered to the patient in order for payment to be made;
- the period of time that the payment is intended to cover, e.g., a day or a month or until the condition being treated is resolved; and
- whether any other services for the treatment of the condition can be billed for separately during the period of time covered by the condition-payment, either

by the provider that is billing for the condition-based payment or other providers providing services to the patient that are related to the condition.

Similar to what is described in subsection 4, a condition-based payment can also be stratified based on patient characteristics by creating a family of condition-based payment codes; this enables a provider to be paid more for managing care of a particular condition if the patient has other characteristics that make care of the condition more expensive.²²⁸

Example: In the "Initiative to Reduce Avoidable Hospitalizations Among Nursing Facilities - Payment Reform," CMS defined six separate G-codes that are intended to pay skilled nursing facilities for care of one of six specific health problems. Each of those codes can be billed for "one day of acute care" for up to 5-7 days following a documented diagnosis, and they can be billed for additional 5-7-day periods after the initial period ends if a clinician documents that the condition is still present. The nursing facility is required to treat the condition in order to bill for the payment, but the payment is not tied to any specific type of treatment service. The nursing facility cannot bill for the code on a day if the patient is treated in a hospital on the same day (either on an inpatient or outpatient basis), even if the nursing facility also provided some kind of treatment to the patient during the day.²²⁹

Since there are thousands of different patient conditions, it would be inefficient to create separate CPT/HCPCS codes for every separate condition, particularly when there are already tens of thousands of ICD-10 codes available.²³⁰ One option is for claims payment systems to evolve so that they could determine payment amounts based on a combination of a CPT/HCPCS code and ICD-10 diagnosis code.²³¹ An alternative would be to use the same condition-based payment code for multiple conditions if it is determined that the appropriate payment amount would be the same. This is similar to the approach used in the Diagnosis Related Groups (DRGs) used to pay for inpatient hospital care – in some cases, one DRG is used for a range of different diagnoses.

Capitation Payments

A special case of the condition-based payment is when the payment is designed to support treatment or management of all or most of the patient's health issues, rather than one condition or a subset of the patient's health problems. For example, "direct primary care practices" charge each of their patients a monthly fee that is not tied to either specific services or specific health conditions. These "capitation" payments can be implemented in the same way as a condition-based payment triggered by a specific condition or combination of conditions, i.e., by creating a billing code that a provider can use to be paid for managing the patient's care.

In those cases, instead of defining the specific conditions which make a patient *eligible* for the payment, it is more likely that there will need to be definitions of those conditions that would justify *excluding* a patient from

the capitation payment and instead paying in a different way, such as through fees for individual services. For example, a primary care practice should not be deterred from caring for a highly complex patient simply because the standard capitation payment is inadequate to compensate the practice for the amount of time that patient will need.

Rather than creating billing codes to enable providers to bill for capitation payments, Medicare and many private health plans have created APMs in which the payer decides whether a primary care practice should receive a "PMPM" (per member per month) payment for a patient based on whether the health plan "attributes" that patient to the practice. This approach not only requires the health plan to create a methodology for attribution, but the primary care practice has to try and predict which patients will be attributed in order to project the practice's revenue and the practice has to try and determine why patients the practice has been caring for were not attributed to the practice. An alternative and more efficient approach is to enable the primary care practice to bill a CPT/HCPCS code for the patient each month. Similarly, instead of a payer trying to risk adjust the PMPM payment based on information derived from claims forms, it would be preferable for the primary care practice to stratify the patients based on characteristics that will require more time or services from the practice. Then the practice could choose a billing code from a family of codes based on the patient's characteristics in order to receive a payment amount that is adequate to cover the time the practice will be dedicating to the patient.²³²

11. Standby Capacity Payments

a. Patient-Specific Standby Payments

As explained under Option 7 in Section VI.A, there are some circumstances in which a standby service is associated with a specific patient who is known to the provider of the standby service, such as a specialist who waits while surgery is being performed on a patient in case the specialist's services are needed immediately for that specific patient. If the specialist's services are needed, they would be paid for those services, but there also needs to be a way to pay for the time they spent waiting if no service is delivered. This can be operationalized in standard billing and claims payment systems by creating a code for the standby time and enabling a provider to bill for it when the service is not actually delivered, using the same approach described earlier for a service delivered in lieu of another service.

Example: CPT code 99360 is used by a physician to bill for their time if they have been asked to be available in case their services are needed and if they have waited to provide services for at least 30 minutes but did not actually perform a procedure. The code can be billed multiple times if the physician remains on standby for 60 minutes or longer.

b. Standby Capacity Payments for Insured Patients

The standby capacity payments discussed at length in Option 7 in Section VI are not associated with specific patients known to the provider; indeed, they are intentionally designed to generate revenue for the provider from “non-patients” who benefit from having the provider able to deliver the service (or having spare service capacity available) but who happen not to need the service during a particular period of time. In some cases, the “non-patient” is a completely healthy individual who needs the standby service in case they are injured or become ill and need immediate treatment.

An individual's health insurance plan is in the best position to determine whether the standby service provider qualifies for a standby capacity payment for an individual and then to make that payment:

- If the standby capacity payment is designed to support a service that equally benefits every individual living in the community, then a health insurance plan should simply make that payment for each of its insured members.

Example: A standby capacity payment from the residents of a community to support an Emergency Department could be paid in this way. The amount of the standby capacity payment would be determined by dividing the ED standby capacity cost by the number of insured residents of the community, and then each payer would pay that amount for each of its insured members.

- If the standby capacity payment is designed to support a service that benefits a subset of patients who have a particular health condition or who are receiving a particular set of services, then a health insurance plan would need to make the payment for each of its insured members who have that condition or who are receiving those services. Since the standby capacity payments are being paid for patients who are part of an APM, the payments could be triggered by other payments made under the APM to care for the eligible condition.

Example: If the standby capacity payment is designed to ensure that a hospital can maintain adequate capacity for labor & delivery services in an APM that is designed to reduce unnecessary C-Sections and increase the use of birth centers, then when a payment under the APM is made for maternity care services of any type to an individual woman, that payment would automatically trigger a standby capacity payment to the hospital.

As explained in Section VI, the standby capacity costs would only represent a portion of the total cost of the services, and the costs associated with actual service delivery would still need to be covered through a fee-for-service payment, a condition-based payment, or one of the other methods described earlier. However, this additional payment would be based on the *marginal* cost of a service rather than the *average* cost. If some users of the service had paid the standby capacity payment and

others had not (e.g., residents of the community vs. visitors to the community who need care from the Emergency Department), then it would likely be appropriate to charge the latter a higher amount for the service than the latter. One way to operationalize the discount for those who had paid the standby capacity payment would be to add a modifier to the service code to indicate that the patient does not qualify for the discounted rate.

c. Methods of Supporting Standby Capacity Other Than Insurance

If health insurance plans are unwilling to make payments for standby capacity, or if a large number of the individuals who would benefit from the standby capacity have no insurance, alternative mechanisms of obtaining standby capacity payments directly from those individuals may be needed. Two options are:

- **Taxation.** A government entity could make the standby capacity payments on behalf of its residents by imposing a tax on the beneficiaries or on their insurance plans. The tax amount could be calculated by dividing the standby capacity cost by the number of taxpayers. In a number of rural communities, the hospital, primary care clinic, nursing facility, etc. are operated by a public hospital district that uses a tax levy to support costs that are not covered by fees for services.
- **Voluntary Contributions.** The individuals who benefit from the capacity could make standby capacity payments through voluntary contributions, similar to the way many small communities use memberships to support their local fire department and EMS services.

12. Volume-Based Adjustments

When there are significant fixed costs associated with a particular type of service and the average cost of delivering the service varies significantly based on volume of services, a volume-based adjustment for that service (Option 8 in Section VI.A) may be appropriate. (The limited number of volume-based payment adjustments that currently exist are made at the facility level, rather than the individual service line, even though a hospital might have a sufficient volume of commonly-used services to deliver them at standard payment amounts but not have sufficient volume to do so for more infrequently-used services.) Moreover, if the volume of services varies over time, it could be the case that a volume-based adjustment is needed at some points in time and not others.

A service-specific volume-based adjustment could be operationalized in at least two different ways:

- by creating a modifier that would be added to one or more service codes to indicate that the volume of services has been or will be below a pre-defined threshold and so a higher amount should be paid for the service;
- by creating a second billing code for a service in addition to the standard code, so that the provider could bill for both the standard payment amount and an incremental payment when the volume of services is

low. If there were additional volume thresholds at which even higher amounts were justified, additional billing codes could be added.

Example: If an APM is paying a primary care practice for delivering care management services to patients with chronic disease, a monthly per-patient payment that is sufficient to enable hiring a nurse care manager at a practice with many chronic disease patients may be insufficient for a practice with few chronic disease patients. In a rural community with only one primary care practice, there would be no way to share the cost of a care manager with other practices, so the patients who do have chronic diseases would only be able to receive the service if a higher amount is paid. A billing code could be created for the monthly payment based on the amount that is appropriate at practices with higher volumes, and then an add-on code could also be created to enable a small-volume practice to receive an additional payment. For example, a low-volume practice would bill both of the following codes for each patient:

- *XXXX01: One month of care management services for a patient with a chronic disease*
- *XXXX02: Additional cost of one month of care management services for a patient with a chronic disease in a practice with fewer than X such patients.*

If the volume of services is variable or if a provider does not know whether the volume of a new service will be high or low, the provider could bill the standard payment when the service is delivered, and then submit a second claim with the add-on code (or a claim with the same code followed by a modifier) after it has been determined that the volume is low. For example, a hospital emergency department could bill for a standard payment as each visit occurs, but then bill for an add-on payment at the end of each month if the total number of visits fell below a certain threshold.

13. Outlier Payments

Outlier payments (Option 9 in Section VI) are needed when the cost that a provider (or group of providers) incurs to deliver a service or group of services to one particular patient is substantially more than what the provider(s) would normally be paid for the service(s). In general, it is impossible to determine the exact difference between cost and payment for any individual service or patient because, as discussed in Section IV.B, the cost of a service for an individual patient depends on the total number of patients receiving the service. Consequently, outlier payments can mitigate shortfalls in payments compared to costs, but they cannot ensure that total payments will exactly match a provider's actual costs.

Three different approaches can be used to operationalize outlier payments, depending on the nature of the gap between costs and payments:

- a. outlier payments for higher-than-expected time or costs;
- b. outlier payments for more services than expected; or

- c. outlier payments for high out-of-pocket costs incurred by the provider.

13a: Outlier Payment for Higher-Than-Expected Time or Costs of Delivering Services

If delivering a service or bundle of services to a particular patient requires an unusually large amount of time or other costs because of the characteristics of the patient, a modifier could be added to the standard billing code for the service or bundle to indicate that a higher payment was needed, or an additional code could be created to trigger an additional payment.

Example: In the Medicare Physician Fee Schedule, Modifier 22 is added to a procedure code to indicate that the work required to provide the service is substantially greater than typically required.

13b: Outlier Payment for More Services Than Expected

If a provider is receiving a bundled payment to deliver a group of services to a patient, and the patient needs an unusually large number of one or more of the individual services that are included in the bundle, the provider could submit a claim with the code for the bundled payment, and also submit codes for all of the individual services that were delivered with modifiers indicating that they were delivered as part of the bundle. The payer could then calculate the total amount that would have been paid for the individual services had they been paid separately, subtract the bundled payment amount, and then pay the provider a proportion of the difference in addition to the bundled payment.

Example: Medicare uses a variation of this approach in the Inpatient Prospective Payment System. Under the IPPS, a hospital ordinarily receives a standard payment amount for a hospitalized patient based on the MS-DRG assigned to the patient, regardless of the exact combination of services that are delivered to the patient. However, the hospital can receive an outlier payment for a patient if the estimated costs of the services delivered to the patient exceeds the standard payment amount for that MS-DRG by a minimum amount. In order to determine the outlier payment, the hospital's charges for each service delivered to the patient are summed and then multiplied by the hospital's cost-to-charge ratio to estimate the actual cost of the case. If that cost is higher than the sum of the MS-DRG payment plus a pre-defined "fixed-loss outlier threshold," Medicare pays the hospital 80% of the difference.²³³

13c: Outlier Payment for High Out-of-Pocket Costs

If a provider has to purchase supplies or medical devices as part of treatment and the cost of a particular product is high and uncontrollable, a supplemental payment can be made to the provider when the provider's out-of-pocket cost for the product exceeds expected amounts.

Example: In the Outpatient Prospective Payment System used by Medicare to pay hospitals for an outpatient service, the hospital receives a pre-determined bundled payment amount for all of the activities directly related to a particular service. However, if specific drugs, biologicals, or medical devices are used in the delivery of the service, the hospital receives an additional "pass-through" payment. The hospital bills for these pass-through items using a HCPCS code beginning with the letter "C" (a "C-Code").²³⁴

B. Determining Eligibility of Patients

As explained in Section VI.D, many current Alternative Payment Models determine which patients are eligible for an APM using methodologies that "attribute" patients to providers retrospectively. Fortunately, the coding and billing mechanisms described above can eliminate the need for these problematic attribution systems. When a provider submits a claim form for a patient using a billing code that is created specifically for the APM, the provider would be explicitly agreeing to take accountability for achieving performance targets for that patient as required under the APM. The provider would know exactly which patients they were responsible for because they had submitted a bill requesting payment for the patient, and the payer would then also know which patients the provider was taking responsibility for. This reduces administrative time and cost for both the payer and the provider.

The provider would be taking accountability for each patient for the specific period of time that is associated with the payment code the provider submits and the accountability methodology that is part of the APM. If the provider is billing for a service bundle or a condition-based payment, the provider would be taking accountability for the time period for which utilization and quality are being measured in the accountability components of the APM. If a patient chose to transfer their care to a different provider, the new provider would bill for the appropriate code, and the payer would know immediately that accountability had shifted to the new provider, rather than waiting for calculations to be made under an attribution methodology.

Two providers that are managing different conditions for the same patient could each take accountability for their portion of the patient's care if the conditions can be managed separately and if there is an APM that gives them the resources to do so. Each provider would submit a claim with a code indicating that they were participating in the APM for the condition they were managing. If the two conditions were so closely interrelated that they needed to be managed jointly, then an APM should be designed to support team care of the two conditions. For example, the two providers could share in a multi-provider bundled payment (Option 12 in Section VI.A) for managing the patient's care. The responsibilities of each provider would be defined by the payment model and their participation would be signaled by the bill they submitted to the patient or payer, rather than having a statistical algorithm assign accountability for everything

to one provider or the other or both. Section VII.A describes how modifiers to billing codes could be used to distinguish the different roles of different providers.

C. Measuring Performance

In addition to defining mechanisms for billing and payment for the services delivered by the healthcare provider as part of the APM, there also need to be mechanisms to operationalize the accountability components of the APM.

1. Measuring Performance on Utilization and Spending

Component #2 of the APM will be based on one or more measures of specific aspects of service utilization and spending that the APM is intended to reduce or control. As discussed in Section VI.B, in many cases these will not be services that would have been delivered by the provider that is being held accountable for reducing/controlling them, so the provider will not be able to directly measure utilization of the services through their own records. Even if the services would have generally been delivered by the accountable provider, the payer will want assurance that utilization or spending on the services by other providers has not increased.

Example: If an APM designed to help primary care physicians improve care of patients with a chronic disease is expected to reduce admissions to the hospital for exacerbations of the disease, the services supported by the APM would be delivered by the primary care physicians but the avoided services would have been delivered by a hospital. Even if a hospital is part of the APM, it is possible that the patient would be admitted to a different hospital, so the accountability component of the APM will need to measure admissions to multiple hospitals.

It would be ideal if the provider that is accountable for performance could directly obtain the data needed to calculate the performance measures so the provider could determine whether the performance targets are being met and if not, to determine why and what to do about it. This would be possible if the provider is participating in a Health Information Exchange (HIE) and all of the other providers who would be likely to deliver services to the provider's patients are also participating. However, there are currently only a few regions in the country where such HIEs exist.²³⁵

In most cases, the only way to obtain the data needed to calculate utilization and spending measures is through the claims data maintained by a patient's health insurance plan, since the plan will have a record of every service the patient received for which a payment was made. If a provider's patients have many different types of insurance coverage, the data for a performance measure will need to come from many different payers, and this will be administratively burdensome for the provider, particularly if every payer stores or reports their claims data in different ways. Some states and regions have established All Payer Claims Databases

(APCDs)²³⁶ that combine claims data from multiple payers, but because of the delays in obtaining and processing these data, measures based on them are generally too old to use for performance-based payments.

a. Addressing the Weaknesses in Claims Data for Calculating Utilization/Spending Measures

Using claims data for utilization and spending measures can be problematic because key information needed for measures is often not available in the data. A good performance measure will focus specifically on the types of services that are avoidable rather than those that are not, and more specifically on the services that could be avoided through actions of the provider(s) participating in the APM. Although CPT/HCPCS codes are recorded in claims data and can be used to determine what *kind* of service a patient received, the same code will generally be used regardless of the *reason* for the service, so current CPT/HCPCS codes are not sufficient to determine whether the service should be included in a measure of avoidable utilization/spending or not.

In some cases, the ICD-10 diagnosis code(s) recorded on the claim can be used to distinguish whether a service should be included in the performance measure for the APM. For example, if a patient with COPD is admitted to the hospital, and the principal diagnosis code for a patient's readmission to the hospital is for trauma-related injuries, it would be inappropriate to count that as a failure to manage the patient's COPD. However, there may or may not be ICD-10 codes for all of the criteria needed to determine whether a service should be included in an accountability measure.

Two approaches could be used to enhance claims data in order to accurately calculate performance measures:

- **Create additional ICD-10 codes.** One or more ICD codes could be created to capture information that is needed to determine whether a particular service should be counted toward the performance measure. For example, there are no ICD-10 codes indicating the stage of cancer, so codes could be created to capture that disease characteristic. However, if multiple criteria are used to determine whether a service should be counted toward a performance measure or not, multiple codes would need to be created, the provider would have to record the codes for each individual criterion, and the payer would need to create an algorithm to determine whether or not the combination of codes for a particular patient met the criteria.
- **Create additional CPT/HCPCS codes or modifiers.** One or more additional CPT/HCPCS codes, or modifiers to existing codes, could be created so the provider of the service can directly indicate whether the service meets all of the criteria requiring it to be counted in the performance measure or not. Different patients might meet the criteria for different reasons, but the same code or modifier would be used unless there was a reason to track the specific criteria that resulted in a particular patient being included or excluded.

Example: In 2005, CMS implemented a Chemotherapy Demonstration Project in which physicians were asked to assess three types of symptoms for patients receiving chemotherapy – nausea/vomiting, pain, and fatigue – on a four-level scale (ranging from "not at all" to "very much"). 12 G-Codes (G9021-G9032) were created (4 for each of the three symptoms to reflect the four different levels on the scale), and physicians were paid \$43.33 for each of the three symptom codes in order to create a total payment of \$130 for symptom assessment (only one of the four symptom level codes could be submitted for each of the three symptoms, and codes for all three symptoms had to be submitted in order to receive payment). The symptom assessment codes could be submitted at each visit when the patient received chemotherapy as indicated by submission of a chemotherapy administration code.²³⁷

Example: In the Deficit Reduction Act of 2005, Congress required that a hospital admission should not be assigned to a Diagnosis Related Group that results in a higher payment solely because of a hospital-acquired infection. In order to distinguish whether a diagnosis code for an infection reflected an infection acquired before or after admission, hospitals are now required to record a "Present on Admission" (POA) code for each Diagnosis Code.²³⁸

b. Insuring Accurate and Complete Coding

For planned services that are being delivered or ordered by a provider or group of providers that is participating in the APM, the default could be to include the service as part of the performance measure unless the diagnosis code, billing code, or modifier is present that indicates it should not be included. The provider would then have a financial incentive to record the codes or modifiers needed to ensure that unavoidable or unrelated services are not included in the performance measure. Of course, there would also be a financial incentive for the provider to inappropriately code avoidable services as unavoidable services, but this is not fundamentally different than the incentive that exists for other types of inaccurate coding, and so the accuracy of coding could be verified through audits or other means.²³⁹

However, unplanned services may be delivered by providers who are not part of the APM, and they would not have any similar incentive to do coding needed to support the APM. This is similar to the problem that exists with diagnosis coding today; diagnosis coding is generally believed to be more complete and accurate for patients admitted to a hospital than for patients receiving outpatient services because the amount that is paid for an inpatient stay can vary depending on the diagnosis codes assigned to the patient, whereas the amount that is paid for an outpatient service does not vary based on the diagnosis code assigned.²⁴⁰ As a result, there is a greater incentive for healthcare providers to ensure complete and accurate coding for inpatient services.

This could be addressed by assigning a higher payment to a service when a provider that is not part of the APM submits the code indicating they had delivered that ser-

vice and also includes the modifier indicating whether the service met the criteria in the APM. The higher payment would reflect the additional time required from the provider to assess whether the service met the criterion or not and to maintain documentation supporting the code or modifier that is chosen. If this enhanced coding were limited to services where there is a potential for overutilization, the higher amounts paid for accurate coding could be more than offset by reducing spending on the services.

Example: In the 2006 Medicare Oncology Demonstration²⁴¹, CMS created 81 new G-codes in order to obtain more detailed data for patients being treated for 13 different types of cancer regarding (1) the stage of cancer, (2) the purpose of oncology visits, and (2) whether the treatment being used adhered to clinical guidelines. Participation was voluntary, and a physician who submitted codes in all three categories was paid \$23 in addition to the standard payment for an Evaluation & Management visit. (The physician received a payment of \$7.67 for the code for each category, but no payment was made unless codes for all three categories were submitted.) There were seven different categories of adherence to guidelines that could be reported using G-Codes:

- *G9056: management of the patient's condition adheres to guidelines*
- *G9057: treatment differs from guidelines due to patient's enrollment in a clinical trial*
- *G9058: physician disagrees with the guideline recommendations*
- *G9059: patient has opted for a different treatment (or no treatment) after being offered guideline-consistent treatment*
- *G9060: patient has comorbidities or performance status not considered in the guidelines*
- *G9061: patient's condition is not covered by guidelines*
- *G9062: management of the patient's condition differs from guidelines for other reasons*

Ideally, health information technology should evolve in a way that would enable this process to be automated. In most cases, the information needed for coding should be recorded in the patient's Electronic Health Record because it is clinically relevant to their treatment, and if that information is stored in a systematic, structured way, software could extract the information and translate it into the appropriate codes for use on claims forms.

2. Measuring Performance on Quality and Outcomes

Component #3 of the APM will be designed to ensure that providers being paid under the APM are maintaining or improving the quality of care, and this will require collecting and analyzing data on one or more measures of quality. As explained in Section VI.C, in some cases it will be preferable to evaluate quality using "process"

measures of quality and in other cases it will be preferable to use "outcome" measures.

a. Obtaining Data on Process Measures and Outcome Measures

One of the reasons that many current process measures are problematic is that they are based solely on the information available on claims forms. In most cases, claims forms only record whether a billable service was performed or not, so it is impossible to determine whether the service was performed in a certain way or whether it achieved a specific goal or outcome, and it is impossible to determine what non-billable services were performed.

However, once a determination is made that data are needed about the way services are performed, what goals/outcomes were achieved, and what non-billable services were performed, it is very feasible to collect that information using existing billing and claims processing systems simply by creating additional codes and recording them on claims forms.

CMS has already created a wide range of what it labels as "Quality-Data Codes" to enable providers to include information on claims forms that describe the quality of care for patients.²⁴² There is a special set of CPT Codes, referred to as "Category II Codes," that are used solely for reporting additional information about services or patients, and CMS has created a number of "G-Codes" and modifiers in the HCPCS system that enable physicians to use claims forms to submit quality measure data for Medicare performance-based payment systems. Some of these codes are pure process measures, i.e., they describe whether a service was performed in a certain way, others are intermediate outcomes, such as the immediate result of a laboratory test or other service, and some are true outcomes, such as a patient's functional limitations. For example:

- **Process Measure:** HCPCS code G9405 is submitted if a patient received follow-up within 7 days following discharge from the hospital
- **Intermediate Outcome Measure:** HCPCS code G9273 is submitted if a patient's blood pressure is below 140/90.
- **Outcome Measure:** HCPCS code G9393 is submitted if a patient with an initial PHQ-9 score greater than 9 (indicating depression) achieves remission after twelve months as demonstrated by a PHQ-9 score of less than five.

In recent years, there has been a focus on obtaining quality data from electronic health records, since much of the information needed to assess the quality of services has already been entered in a patient's medical record. However, payers do not and should not have direct access to patients' medical records, so providers still have to extract the necessary data from the EHR. Although this could be done by creating a quality reporting system that is separate from the claims payment system²⁴³, it is not clear that this has any significant advantages over taking the same information from medical records and submitting it through claims payment systems using CPT Category II and HCPCS codes that

receiving physical therapy to improve their mobility, G8978-CM would indicate that at the beginning of therapy, the patient only had between 1%-20% of full mobility, G8979-CJ would indicate that the goal of therapy was to achieve 60%-80% mobility, and G8980-CI would indicate that at the completion of therapy, the patient had actually achieved better than 80% of full mobility.²⁴⁵

The feasibility and desirability of this approach depends on the extent to which small differences in quality in each category can be treated as equivalent. If there is an important difference in the impact on a patient when a metric value is at one end of a category or the other, then further disaggregation may be desirable so the ranges for the categories are smaller. On the other hand, if there is significant measurement error involved with the metric, then wide categories may actually be a more robust way of assessing performance than trying to reward or penalize providers for small differences in the measure.

d. Coding for Comparison Populations

As noted in Section VI, one of the difficulties of assessing performance with respect to comparison groups is that the data needed to define the comparison group or to measure performance on the group will not be available if those data are not currently collected or submitted by providers who are not participating in the APM. In contrast to providers participating in the APM, there might be no reason for other providers to incur the costs of collecting and submitting additional data. This could be addressed by allowing other providers to be paid for submitting the codes with the necessary data.

D. Adjusting Payments for Performance

Section VII described several different options for adjusting a provider's payments based on the provider's performance on utilization/spending and quality/outcome. Each of these would have to be operationalized somewhat differently.

1. Penalties/Bonuses Based on Achieving Targets

a. Patient-Level Performance Targets

The easiest thing to implement in a standard billing and claims payment system is a bonus payment based on achieving a patient-level performance Target. In addition to whatever billing codes have been used to obtain payment for the services that have been delivered, the billing code or modifier that is associated with the performance measure would be submitted when the provider achieves the Target for a patient, and the provider would then receive an additional payment for that patient. The sum of all of those additional payments would be the total performance-based bonus the provider receives.

Example: In the Medicare Diabetes Prevention Program (MDPP), a provider can bill one of three HCPCS G-Codes (G9873-G9875) for an eligible beneficiary who attends core sessions during the first 6 months in which they are enrolled in the program and receive payments ranging from \$25-\$90 depending on the number of sessions attended. If the beneficiary achieves a 5% reduction in their weight during this period, the provider can submit a bill with HCPCS code G9880 and receive an additional payment of \$160 and if the beneficiary achieves a 9% reduction in weight, the provider can also submit a claim with code G9881 and receive a second additional payment of \$25.²⁴⁶

This approach can also be used to implement a penalty for failure to achieve a patient-level performance Target simply by (1) decreasing the standard amount that is paid for the services (whether the payment is through fees for individual services, a bundled payment, a condition-based payment, etc.) and then (2) paying a bonus for each patient for whom the performance Target is reached. The amount of the reduction in the payment for services would be such that when the provider achieved the minimum performance level needed to avoid a penalty, the sum of the bonus payments would be equal to the sum of the reductions in the payments for the services/conditions, i.e., the provider would receive the same amount of revenue as if there was no performance adjustment. This is equivalent to what is commonly described as a "withhold" – a portion of the provider's payment is withheld and paid only after the necessary performance has been achieved.

Use of a performance-based billing code facilitates the use of outcome measures that can only be measured after a delay. The provider can bill and be paid for the service when it is delivered (albeit at a lower amount), and then bill and be paid for the performance-based payment as soon as the outcome is achieved. If the outcome is not achieved, this creates a penalty since the payment is less than it would otherwise have been.

This approach is preferable to the way that most current penalty/bonus payments in both APMs and standard fee-for-service payments are implemented. In typical retrospective, payer-administered systems, the payer pays unadjusted amounts for services as they are delivered, collects information on utilization/spending and/or quality/outcomes after the services are delivered, calculates the provider's performance, and then pays a bonus or imposes a penalty on the provider months or even years after the services were delivered. This approach has a number of significant weaknesses:

- The penalties or bonuses are paid long after the services were delivered. A provider could be forced to pay a penalty in a particular year based on what happened in a prior year, even though the individuals who caused the problem are no longer working there.
- It is difficult, if not impossible, to reduce the payments already made by individual patients when penalties are owed to them or to collect additional payments from patients when bonuses are owed by them.

are designed for that purpose. Moreover, it is difficult to determine performance-based payments using both utilization/spending measures and quality measures if those measures are being reported through two different systems. As discussed in the previous section, HIT systems could be developed to automatically extract the relevant data from a properly-designed EHR and translate them into the appropriate codes for submission on claims forms, thereby minimizing administrative burden on providers.

If the data needed for a quality measure are not currently being collected, but it would be feasible for the provider to collect the data, the provider could use whatever method for data collection is most feasible and report the results through standard billing and claims data systems using codes designed for that purpose.

An advantage of using special codes in claims payment systems for quality measurement is that it easily allows a provider submitting the code to be paid solely for the collection and submission of the data if there is a significant cost associated with that or if the data are coming from a provider or other entity that is not directly participating in the APM and would otherwise have no reason to collect or report the information. One of the barriers to collecting data on outcomes through a patient registry is that there is currently no direct way for providers to support the cost of maintaining the registry. However, if data from the registry were submitted to a payer through billing codes, the payer could pay for those data and thereby help support the cost of maintaining the registry.

Example: In the 2006 Medicare Oncology Demonstration, CMS created 81 new G-codes in order to obtain more detailed data for patients being treated for 13 different types of cancer regarding (1) the stage of cancer, (2) the purpose of oncology visits, and (2) whether the treatment being used adhered to clinical guidelines. Participation was voluntary, and a physician who submitted codes in all three categories was paid \$23 in addition to the standard payment for an Evaluation & Management visit. (The physician received a payment of \$7.67 for the code for each category, but no payment was made unless codes for all three categories were submitted.)²⁴⁴

b. Collecting Data Through Other Mechanisms

The process described above can be used for any quality measure, including outcome measures, as long as the necessary data can be collected by the provider participating in the APM. However, there are a variety of circumstances in which alternative methods of collecting the data are necessary or desirable, such as:

- It may be impractical for providers to collect the data because of distance or other factors. For example, if a patient traveled a long distance from home to obtain a particular procedure from a specialist, it will not be feasible for the specialist to directly evaluate the patient's condition after the patient returns home.
- If the measure in question requires a subjective assessment of the patient's condition, there may be

concerns that the provider's judgment will be biased in favor of a positive assessment.

- If the measure in question is based on a patient survey, there may be concerns that the response rate will be biased toward the patients that the provider can contact easily, toward the patients who will respond in a particular way, etc.

In these cases, a separate mechanism for collecting the data may have to be established, which will in turn require a way of paying for the cost of that data collection mechanism. The best approach to doing this will likely depend on the circumstances. In some cases, it may be preferable for one or more payers to establish (and pay for) the data collection mechanism; in others, it may be preferable for an independent entity to operate the data collection mechanism and to contract with payers or providers or both in order to obtain the funds needed.

c. Coding of Continuous Quality Measures

Many quality metrics are inherently continuous rather than categorical. Even if there is a threshold level on the quality metric that is viewed as desirable, a patient may still receive some benefit if the care delivered falls just short of that threshold. If a smaller shortfall is more desirable than a bigger shortfall, the accountability component would need to differentiate between those situations.

It is impossible to precisely translate continuous quality metrics into billing or diagnosis codes because the codes are inherently categorical; one would need an extremely large number of codes to represent all of the many different values of the metric. As a result, continuous metrics and measures are typically converted into categories using ranges of the different values the metric or measure can take on.

Once categories for the metric have been defined, a code can be defined for each of the categories. If the level of quality achieved on the metric falls within a particular category, then the code for that category is recorded and submitted.

Example: In 2012, Congress required that CMS implement a claims-based system for reporting on improvements in patient functional limitations resulting from physical therapy, occupational therapy, and speech-language pathology services. Under the system CMS established, therapists report data on patient functional limitations using 42 G-Codes and 7 severity/complexity modifiers. There are 14 sets of G-codes for 14 different types of functional limitations (e.g., mobility, swallowing, speech, etc.), and there are 3 codes in each set – one is used to report the level of the patient's functional limitations when therapy begins, one is used to report the goal for the level of functional limitation that the therapy is intended to achieve, and one is used to report the level of functional limitation when therapy ends. The 7 modifiers are used with each G-Code to indicate the percentage of impairment on the specific type of functional limitation indicated by the G-Code; each modifier specifies a range of percentages, e.g., 1-20%, 20-40%, etc. For example, if a patient was

b. Population-Level Performance Targets

Bonuses/penalties for population-level Targets cannot be implemented in this way if the penalties or bonuses are not directly proportional to the number of patients or services. In these cases, it may still make sense for providers to submit a code when success has been achieved for an individual patient, but these codes would provide a way of measuring how often success was achieved and for which patients it was achieved, rather than being used to directly trigger a payment. The provider and payer would then have to make a determination at the end of the performance period (or after a minimum number of patients had been treated) if the Target performance level had been achieved, and what the appropriate penalty or bonus should be.

2. Outcome-Based Payments

Option 2 in Sections VI.B and VI.C requires the use of a patient-level Target, but since the payment for the service or patient is contingent on achieving the Target, there would not be a need to create a separate code to signal that fact. The provider would only submit a claim for payment if the Target had been achieved.

Example: In the Medicare Diabetes Prevention Program, a participating provider only receives payments for delivering services to a patient during the patient's second year of participation in the program if the patients achieve or maintain a 5% reduction from their baseline weight when they entered the program. If a patient does not achieve or maintain the 5% weight loss, the provider receives no payment for the services delivered to that patient.²⁴⁷

For outcomes that can only be measured after a long period of time, it may be desirable for the provider to receive a partial payment when the service is delivered, and then receive the balance of the payment when the outcome is achieved. In this situation, it would still make sense to create two separate codes, with one signaling that the service was delivered so the partial payment can be made and the second code signaling that the target was achieved and the balance should be paid. If the payments are being made by an insurance plan with multiple patients who received services, then the payment to the provider could be the net amount of the additional payments for patients who achieved the outcomes minus recoupments of the partial payments for patients who did not. Otherwise, the provider would need to refund the partial payment to an individual patient or their payer if the outcome was not achieved.

3. Bundled/Warrantied Payments

Under Option 3 in Sections VI.B and VI.C, if an avoidable service was delivered, or if an additional service was needed to correct a defect in quality (or if some form of compensation were paid for the defect), the accountable provider would be required to pay for that from the bundled/warrantied payment. The same mechanism for implementing a bundled payment for planned and unplanned services described earlier could be used to implement this approach.

VIII. IMPLEMENTING THE ALTERNATIVE PAYMENT MODEL

STEP 6 Implement the APM, assess its performance, & make improvements

A. Obtain Participation by Payers, Providers, and Patients

- Take actions to encourage payer participation
- Design APMs to encourage provider participation
- Design APMs to encourage patient participation

B. Finalize the APM Parameters

- Use a beta-testing process for innovative APMs

C. Evaluate the APM

- Define what should be evaluated
- Define the timeframe for evaluation
- Set thresholds for success
- Take actions based on the evaluation results
 - ◆ Expand use of the APM
 - ◆ Modify the APM
 - ◆ Terminate the APM
 - ◆ Continue payment model through other means

D. Revise/Update the APM Parameters

No matter how well an APM is designed, it will not be successful unless it is actually implemented. Moreover, its total impact will depend on how broadly it is ultimately used, and whether it is appropriately adjusted over time to address unanticipated problems and to adapt to changing circumstances. There are four separate steps needed for success:

- Obtaining agreements by payers, providers, and patients to participate in the APM;
- Finalizing the details of the APM design;
- Evaluating the APM to make decisions about continuation/expansion; and
- Updating the APM parameters over time.

A. Obtaining Participation by Payers, Providers, and Patients

An APM is nothing more than a concept until at least one payer agrees to implement the APM, at least one provider who is paid by that payer agrees to participate, and at least some of the patients insured by the participating payer and receiving care from the participating provider are willing to accept the approach to care delivery supported by the APM.

1. Encouraging Participation by Payers

a. Why Payers Don't Implement APMs

Unfortunately, many payers have failed to implement APMs even when there are significant opportunities for savings or quality improvement and there are documented barriers in the current payment system that prevent those opportunities from being achieved. There are several common reasons for this:

- Administrative costs for payers to implement the APM;
- Disincentives for insurance companies to encourage reductions in healthcare spending;

- Benefits to individual payers of being a “free rider;” and
- Barriers in provider contracts.

i. Administrative Costs for Payers

A payer will have to incur a variety of costs to implement an APM, such as revising the programming in its claims payment system to utilize new billing codes, modifying contracts with providers to include new accountability mechanisms, and potentially revising the rules governing the benefits for patients and seeking approval from regulatory agencies to do so. The fact that an APM will reduce healthcare spending does not necessarily mean there will be a business case for a payer to implement it if the administrative costs are higher than the expected savings.

It is not surprising that the most common “payment reforms” being implemented by payers are pay-for-performance programs and shared savings models because these approaches require no changes in the systems payers use to pay providers for the individual ser-

vices they deliver; they merely require the payer to calculate spending and quality measures, often just once per year, and to make one-time bonus payments or penalties.

ii. Disincentives for Participation by Insurance Companies

Unfortunately, even if the healthcare savings created by an APM would be significantly greater than the administrative costs of implementing the APM, there will not be a business case for most health insurance companies to implement it because they are not the true purchasers of the care, i.e., the money to pay for healthcare services is ultimately coming from someone else. A payer that is not the ultimate purchaser will incur the administrative costs of implementing an APM but retain little or none of the savings in healthcare spending that the APM helps to create. Moreover, over time, insurance company profits are generally higher when healthcare spending is higher, so even if administrative costs associated with an APM are low, there will be a disincentive for the insurance company to implement APMs that reduce healthcare spending.

The reasons for this differ for different health plans in different communities:

- **ASO Contracts with Self-Insured Purchasers.** The majority of individuals “insured” by most commercial health insurance companies are covered under what are known as “Administrative Services Only (ASO)” contracts, where an employer, retirement plan, or other purchaser hires the insurance company to pay claims on behalf of the purchaser’s employees or members.²⁴⁸ The purchaser is “self-insured” and accepts all of the financial risk for healthcare spending, and the purchaser pays the insurance company a fee for processing and paying claims and for providing other services to the purchaser. Under ASO contracts, the costs of implementing an APM will be incurred by the insurance company (since it actually makes the payments to healthcare providers), but any savings from lower healthcare spending will go to the purchaser, not to the insurance company. As a result, implementing the APM will represent a financial loss to the insurance company unless it can charge higher administrative fees to the purchaser. In addition, if administrative fees are based on the volume or cost of claims, then a reduction in utilization or spending will reduce profits for the insurance company.

Insurance companies with ASO contracts also have difficulties implementing APMs that use population-level accountability measures and/or retrospective reconciliation payment methods. Under fee-for-service payment systems, the purchaser knows that any payments it made were directly associated with the employees or members it insures. However, under payment models where the same payment is made for every patient, an *individual* purchaser might be spending more than they would have under a fee-for-service system even if *aggregate* spending for *all* purchasers is lower. If performance measures for spending are calculated at the population level rather than the patient level, when savings occur for some

patients, those savings will be credited to all purchasers. For the subset of patients who actually had lower spending, their purchaser may pay more than if they had paid fees for individual services. For example, in a typical shared savings APM, the benchmarks and savings calculations are based on averages across all of the provider’s patients. In order to make a shared savings payment to the provider, the insurance company would have to charge a portion of that payment to the self-insured purchasers, but there is generally no reliable way of determining what portion of the savings is attributable to each individual purchaser.

- **Insurance Products.** For individuals covered by true insurance plans managed by commercial insurance companies (i.e., where the insurance company accepts premiums from individuals or from entities that are purchasing insurance on an individual’s behalf, in return for a promise to pay the healthcare expenses for those individuals from the collective premiums), the health insurance company would receive the healthcare savings the APM would produce as well as incur the costs of implementing the APM. However, federal law now requires that commercial health insurance plans spend 80-85% (i.e., the “minimum medical loss ratio”) of their premium revenues on healthcare services.²⁴⁹ That means that if there is a reduction in healthcare spending, the insurance company would be required to reduce its administrative costs or profits. If administrative costs have increased due to implementation of the APM, the company’s profit would need to decrease even more, which creates a disincentive for a profit-making insurance business to implement the APM.
- **Capitated Plans.** Medicare Advantage plans (MA plans) and Medicaid Managed Care Organizations (MCOs) do not accept full insurance risk in the same way that a commercial health plan does. They receive a risk-adjusted payment for each enrolled member from CMS or a state Medicaid agency, which means that CMS or the state Medicaid agency remain a partially self-insured purchaser. In the short run, MA plans and MCOs would retain any savings generated by an APM, and that could offset the administrative costs they incur by implementing the APM. However, if their patients are healthier, the patients’ risk scores would decrease; this would reduce the risk-adjusted payments the plans receive, thereby reducing the plans’ revenues while leaving them with higher administrative costs.

iii. Benefits to Payers of Being a Free Rider

In addition to the above issues, in any community where there are multiple payers, an individual payer will likely benefit more if it does not implement the APM while other payers do. If a provider’s participation in the APM results in changes in the way the provider delivers care to patients, the provider will generally make the same changes for all of their patients, not just those whose insurance plans are using the APM. In most cases, it is the change in care delivery that produces savings, not the APM *per se*. Consequently, a “free rider” – a payer that is not participating in the APM – will receive the savings from the changes in care to its patients without

having to incur the administrative costs of implementing the APM or to make additional payments to the provider to support the costs.

iv. Barriers in Provider Contracts

Some large providers refuse to contract with payers unless the payers agree to contract provisions that limit or prevent the payer from taking actions that would be required for implementation of an APM. For example, "anti-tiering" and "anti-steering" clauses prevent payers from charging patients higher cost-sharing amounts when they obtain services from higher-cost providers.²⁵⁰

b. How to Encourage Payer Participation in APMs

There are several ways of encouraging payer participation in well-designed APMs. One or more of these approaches will likely need to be part of a strategy for implementing an APM in a particular community.

i. General Approaches to Encourage Payer Participation in APMs

- **Design the APM to work within existing payer administrative systems.** An APM will be more attractive to payers if the administrative costs associated with implementing it are as low as possible. An APM that is designed to operate within existing claims payment systems using the approaches described in Section VII will likely require significantly lower administrative costs than an APM that requires development of entirely new systems. (In addition, an APM that operates within existing payer claims payment systems will also likely be easier and cheaper to implement within provider billing systems, which will help to encourage participation by providers.)
- **Use a common approach to coding and definitions in APMs.** There is currently a confusing proliferation of G-codes, S-codes, and Medicaid codes used in APMs. Administrative costs are higher for both payers and providers if every APM uses different codes to represent the same concepts, and costs for providers are higher if different payers use different codes for the same concept or use different definitions of the same code. A mechanism for developing agreements on coding and definitions for APMs is needed, similar to the processes that enable the use of the same CPT and ICD codes for procedures and diagnoses under the current fee-for-service system.²⁵¹
- **Use APMs designed in ways that can be used with self-insured purchasers.** APMs that use condition-based payments, bundled payments stratified by patient characteristics, and patient-specific performance targets are more likely to show a positive business case for self-insured purchasers, particularly those with small numbers of employees/members, than APMs using shared savings, capitated payments, or population-level accountability measures.
- **Require payers to publicly disclose the payment methods they use.** In order for purchasers and patients to selectively choose payers that use APMs, they have to know what methods a payer uses to pay providers. Today, it is virtually impossible to find out how a payer

pays providers. If payers were required to disclose the methods by which providers are paid (or if providers were not prohibited from disclosing this information by provisions in their contracts with payers), patients and purchasers could choose the payers that use APMs that will result in lower spending and higher quality.

Although more information on payment systems has become available in recent years than in the past, unfortunately, the information in these reports is incomplete and can be misleading or harmful. In particular:

- ◆ "Price transparency" initiatives tend to focus on the *amounts providers are paid for individual services*, not the *method by which the provider is paid*. This can inadvertently penalize payers and providers participating in APMs in which the payment for a particular service is higher because it is being used less frequently, is being substituted for more expensive services, or has a lower rate of complications. If a provider that is paid a lower price for a service has a higher rate of complications but the payment structure requires no accountability for that, choosing the lower-priced provider could result in worse outcomes for the patient and higher total spending on their care.
- ◆ Reports on the extent to which payers are using "value-based payment" systems are generally based on simplistic categories that treat the use of small bonuses or penalties based on population-level process measures as equivalent to patient-centered, outcome-based payments using patient-level measures. The reports make no effort to determine whether the payment system being used adequately addresses the barriers in current payment systems that are preventing the delivery of higher-value care.²⁵² Moreover, these reports typically allow payers to self-classify their payment systems with no documentation on the details of the payment system or external validation of the classifications. This can imply that more progress is being made in using well-designed APMs than is actually occurring.
- **Prohibit provisions of payer-provider contracts that limit the ability to implement desirable APMs.** Federal or state legislative or regulatory actions could be taken to prohibit provisions in the contracts between payers and providers that make it difficult for payers to implement APMs.

ii. Purchaser Actions to Encourage Implementation of APMs

Purchasers (including individuals purchasing insurance on exchanges) are ultimately those who will spend more if APMs are not implemented, and there are additional actions they can take to encourage implementation of well-designed APMs:

- **Select payers based on willingness to implement APMs.** Under the business model for most insurance companies, lower spending per member results in lower profits, which means the payers' interests are not aligned with those of the purchasers (e.g., self-insured employers and individuals purchasing an individual insurance policy on an exchange) who would benefit from lower spending per member. However, if

the number of insured members increases, then even with lower spending per member, a payer’s total spending and profits can increase. Consequently, if employers, individuals, and other purchasers only use payers that implement a desirable APM, there will be a stronger incentive for payers to implement that APM.

- **Contract for insurance and care delivery through purchaser coalitions.** The willingness and ability of a payer or provider to participate in an APM will depend on how many patients will be participating in the APM. There will inherently be some administrative costs involved in implementing an APM, and the cost per patient will be lower if there are more patients participating. In addition, the threat of losing members due to failure to implement an APM will be more powerful if a larger number of members are involved. Small purchasers can have a bigger impact by working together through purchaser coalitions than by trying to influence payers individually.
- **Use direct purchaser-provider contracting.** The ability for a purchaser to selectively choose a payer that implements an APM depends on there being at least one payer that is willing to do so. In markets where there is only one payer, or where none of the payers are willing to implement an APM in order to attract more members, purchasers could choose to become the payers themselves, contracting directly with the providers and cutting out the middle-man. A growing number of self-insured employers are doing this.²⁵³ APMs that are designed to work for self-insured purchasers will facilitate direct purchaser-provider contracting.

Purchasers who want to encourage successful use of APMs need to make a commitment to support them for multiple years. As discussed further below, providers are unlikely to make major changes in the way they de-

liver care in response to a payment reform that may only last for a year or two. Purchaser-led initiatives in the past have failed because purchasers changed to payers that offered lower short-run premiums even though those payers were not committed to long-run reforms in payment.²⁵⁴

iii. Provider Actions to Encourage Implementation of APMs

Finally, there are actions that at least some providers can take to encourage payer participation in APMs.

- **Refuse to contract with payers that do not implement APMs.** If at least one health plan is willing to implement an APM, providers with sufficient market power could refuse to contract with other health plans that do not implement the APM. If this causes the other payers to fall short of network adequacy requirements under insurance regulations, the other payers would either have to implement the APM or exit the market.
- **Develop the capability to contract directly with purchasers or to sell insurance products.** A number of providers, including associations of small physician practices, have created their own health insurance plans in order to be able to contract directly with purchasers and so they can be paid in the right way.²⁵⁵

Reasons Why Payers Don't Implement APMs	Ways to Encourage Payer Participation in APMs
<ul style="list-style-type: none"> • Payers incur administrative costs to implement APMs that weaken the business case for the APM • Reducing healthcare spending reduces profits for health insurance companies • Payers benefit by being a “free rider” while other payers implement an APM • Contracts with providers prohibit implementation of APMs 	<p>APM Design</p> <ul style="list-style-type: none"> • Design APMs to work within existing payer administrative systems • Use a common approach to coding and definitions in APMs • Design APMs that work for self-insured purchasers <p>Government Actions</p> <ul style="list-style-type: none"> • Require payers to publicly disclose the payment methods they use • Prohibit provisions of payer-provider contracts that limit the ability to implement APMs <p>Purchaser Actions</p> <ul style="list-style-type: none"> • Select insurance plans based on willingness to implement APMs • Contract for insurance and care delivery through purchaser coalitions • Use direct purchaser-provider contracting <p>Provider Actions</p> <ul style="list-style-type: none"> • Refuse to contract with payers that do not implement APMs • Develop the capability to contract directly with purchasers or to sell insurance products

2. Encouraging Participation by Providers

There is little reason for payers to implement an APM or for purchasers to encourage them to do so if it is unlikely that healthcare providers will want to participate in the APM. The fewer providers who participate in the APM, the smaller the number of patients who will receive the care the APM was intended to enable. This will lead to a smaller amount of total savings, and the administrative costs payers will incur will be higher on a per-patient basis.

a. Why Healthcare Providers Don't Participate in APMs

Lower-than-expected participation in APMs is often attributed to a preference by providers for traditional fee-for-service payment. However, in most cases, there are other reasons that providers don't want to participate in APMs²⁵⁶, including:

- problems with the design of the APM;
- the small number of payers who would be using the APM, or inconsistencies among payers in the structure of the APM or the measures used in the APM;
- the inability to cover extra costs incurred during the transition to the APM;
- lack of reserves to manage financial risk required under the APM;
- lack of data to estimate potential savings and risks under the APM;
- no assurance of stability or continuation of the APM;
- failure of the APM to address specific types of patient needs or unique issues in the community;
- requirements in federal or state laws or regulations that prohibit or limit the ability to implement the APM; or
- unwillingness of the provider to make the reductions in cost or improvements in quality needed to succeed under the APM.

i. Problems with APM Designs

Although there are many ways in which an APM can be designed badly, the problems that most commonly discourage provider participation include:

- **Failure to remove the barriers in fee-for-service payment.** Most current APMs do not remove the barriers to changing care delivery that exist in the current payment system, making it difficult or impossible for providers to achieve savings or improvements in quality.
- **Penalties for things outside the provider's control.** Many current APMs hold providers accountable for controlling total spending on healthcare services for their patients, even though the providers cannot control many of the factors affecting total spending.
- **Retrospective statistical attribution of patients.** Most current APMs use statistical algorithms to determine the patients for whom a provider will be held accountable for spending and quality. Since these algorithms are applied after services have already been deliv-

ered, it is impossible for the provider to change care delivery for those patients. Moreover, since most algorithms only use information available from claims forms, patients can be incorrectly assigned to a provider who was not actually responsible for managing their care.

- **Retrospective performance targets.** Many current APMs define the targets for spending or quality after services have already been delivered, making it impossible for a provider to know in advance how much they will be paid or what performance standards they should try to meet.
- **Inadequate adjustment for differences in patient needs.** Many current APMs do not adjust payments for differences in patient needs or adjust accountability measures for differences in patient risk, which can cause providers to be penalized when they deliver appropriate treatments to higher-need patients. Those APMs that do adjust payments often use risk-adjustment systems that fail to consider some of the most important characteristics of patients affecting the services they need and the outcomes that can be achieved.²⁵⁷
- **Excessive financial risk.** Many APMs require that when providers fail to reduce a payer's spending, they must pay a penalty based on a percentage of the payer's spending. However, if the provider's revenue is only a small fraction of the payer's spending, a "small" penalty from the payer's perspective could represent a large fraction of the provider's revenue, placing them at risk of bankruptcy.
- **Burdensome administrative requirements.** Many APMs require providers to collect and report large numbers of quality measures. These requirements cause providers to incur costs that they may not be able to recoup through the APM itself.
- **Infeasibility for small providers and rural communities.** Many APMs to date have been designed in ways that explicitly or implicitly favor large provider organizations and integrated delivery systems, making it difficult or impossible for small physician practices, small hospitals, and other small providers to participate.

ii. Lack of Multi-Payer Participation

Even if an APM is well-designed, a healthcare provider may still be reluctant to participate if only one payer is participating in the APM while other payers continue to use standard payment systems. There are two reasons for this:

- **Inability to use a single approach to care.** Because the APM is designed to support a different approach to care delivery that cannot be sustained under the standard payment system, the provider would presumably only be able to financially support the new approach for those patients whose payers participate in the APM. However, no healthcare provider will want to deliver less effective care to a subset of their patients based on the type of health insurance they have. Moreover, using two different approaches to

care delivery for the same health conditions creates inefficiencies and higher costs for the provider.

- **Higher average cost of new or changed services.** If there are significant fixed costs associated with implementing the new or changed services, the average cost of the services will be higher if fewer patients are receiving them. As a result, the payment amounts under the APM may be inadequate to cover the provider's costs if the payments are being made only for a subset of the eligible patients.

CMS has recognized that even though Medicare is the payer for a large percentage of physicians' and hospitals' patients, it rarely represents the majority of a provider's patients and it may represent an even smaller proportion of the provider's revenues, so it has attempted to encourage multi-payer participation in Medicare APMs. However, the methods it has used to date have been problematic. In the Comprehensive Primary Care Initiative and Comprehensive Primary Care Plus, CMS did not permit primary care practices to participate unless they were located in states or substate regions where a sufficient number of other payers also agreed to pay the practices in similar way. Although the goal was laudable – enabling participating primary care practices to receive the majority of their revenues through the APM – the approach meant that payers who did not want to implement the APM had veto power over whether primary care practices in their communities could be paid differently under Medicare. Moreover, while CMS has encouraged other payers to participate in CMS-designed APMs, CMS has generally been unwilling to participate in APMs designed by other payers.

In addition, even when multiple payers are nominally implementing the “same” APM, there are often differences in payment amounts, coding requirements, quality measures, risk adjustment methods, etc. For example, many different payers have implemented “bundled payments for joint surgery,” but the definitions of what is included in the bundles differ significantly. At a minimum, these differences likely increase administrative costs for providers, but in some cases, the differences could cause a provider to be rewarded under one version of the APM for delivering care in a certain way while being penalized for the same thing under another version.

iii. Inability to Cover Transition Costs

In most cases, the bigger the potential savings, the more complex the changes in care delivery that will be required to achieve them, the longer it will take to implement those changes, and the greater the costs a provider will likely incur in making the transition.

This means that even if the APM is designed well, and even if the amounts of payment under the APM would be adequate to cover a provider's costs once the transition to the new mode of care delivery has been completed, payments under the APM may not be adequate to cover the costs the provider incurs during the transition phase, and the APM may not produce net savings.

In other industries, a business that develops a new product or service will generally expect to lose money on the product or service initially. The business will cover

the short-term losses either by spending retained earnings it has accumulated from previous products/services, by borrowing from lenders, or by obtaining funds from outside investors, and then it will repay those funds through profits on future sales, particularly if the volume of sales increases. However, these strategies are less likely to be available for healthcare providers for several reasons:

- **Lack of retained earnings and capital reserves.** Most physician practices do not have retained earnings that can be used as working capital, and small hospitals and other small providers are unlikely to have significant capital reserves.
- **Savings go to payers, not providers.** In other industries, success comes from delivering the product or service to additional customers, and the return on the investment in a new product/service comes directly to the business through higher revenues from additional customers that can be used to repay debts. However, in most APMs, the goal is to deliver different services to the same patients, not to increase the number of patients receiving services, and to reduce spending rather than increase revenue. The return on investment comes to payers, not the provider, so the provider has no ability to use the savings to repay a loan or investment.
- **Inability to set prices.** In Medicare, payments are set by CMS, so healthcare providers do not have the ability to set higher prices for services in order to recoup borrowing costs, and in the private sector, small healthcare providers may not be able to negotiate higher prices from large payers.

iv. Lack of Reserves to Manage Risk

The lack of retained earnings not only means that providers cannot cover transition costs, it also means they do not have the financial reserves needed to manage significant financial risk. Under the fee-for-service system, a provider is only at risk for losses in revenue due to reductions in the services it delivers itself, but under many APMs, the provider could lose revenue or have to pay penalties if other providers deliver more services or if pharmaceutical companies or other providers raise their prices, and those losses could exceed the total amount the provider receives for its own services.

The premiums charged by health insurance companies are required to be higher than the amount they expect to spend in order to build up reserves. However, in APMs that are explicitly intended to transfer financial risk to providers, there has been no provision for enabling providers to build financial reserves or to transfer the payers' reserves to the providers along with the transfer of risk.

v. Lack of Data to Estimate Potential Savings and Risks

If an APM requires a provider to take accountability for reducing utilization or spending on services that it does not deliver itself, the provider needs data on how often those services are currently being delivered and how much they cost in order to estimate the magnitude of

the potential savings that could be achieved and the probabilities of financial gains and losses under the APM. If the provider agrees to participate in the APM, it will need to receive those data on a timely, ongoing basis in order to determine whether actual results are deviating from expectations and to make changes in the way it is delivering care.

Most providers do not have access to a health information exchange or healthcare claims data that would enable them to determine what kinds of services their patients are receiving from other providers and how much is being spent on those services. When they do receive reports from claims data, they are often more than a year old, making them difficult to rely on for planning purposes and impossible to use for real-time management.

vi. No Assurance of Stability or Continuation of the APM

Most Alternative Payment Models in Medicare are explicitly structured as time-limited demonstration projects, with no assurance of continuation even if the APM is successful in achieving savings or improving quality. Moreover, CMS can and does change the structure and standards in the APM during the course of the demonstration without the consent of the participants. In the private sector, there is no assurance that an APM will continue past the end of the provider's current contract with the payer or that the payment amounts or number of patients insured by the payer won't decrease during the course of the contract.

Physicians, hospitals, and other healthcare providers are unlikely to be willing to fundamentally change the way they deliver care if there is uncertainty about how long the payment changes needed to support the care changes will be in place. The bigger the changes in care delivery, the longer it will likely take to fully implement the changes, and the longer it will take to reverse them if there is no longer a payment model to support them. These delays will create losses during the initial implementation period without adequate time to recoup them, and they will create additional losses after the APM is terminated. Similarly, if changes are made in the structure or parameters of the APM, the changes in care delivery may no longer be financially viable.

vii. Failure to Address Local Differences in Patient Needs and Provider Capacity

The health needs of patients differ from community to community based on factors ranging from the types of jobs in the local economy to the weather. In addition, healthcare services are delivered differently in different communities, for reasons as varied as population density, the ability of the community to attract and retain healthcare workers, and accidents of history. Differences in health needs and in the structure of healthcare services often translate into differences in the opportunities for reducing spending and/or differences in the services needed to address those differences.

viii. Statutory or Regulatory Barriers

Several different laws and regulations have been created at both the federal and state levels to try and prevent specific kinds of problems and abuses that have occurred or are feared to be possible under the current fee-for-service system. Even if an APM includes a better mechanism for preventing the same type of problem or abuse, these laws or regulations will still be in effect and the actions taken by a provider or payer to implement the APM may be technically illegal or at least restricted in some way that makes implementation more challenging. For example:

- **Prohibitions on physician referrals of patients to entities with which they have a financial relationship.** The federal Ethics in Patient Referrals Act, commonly known as the "Stark Law,"²⁵⁸ prohibits physicians from referring Medicare and Medicaid patients to entities such as hospitals with which the physicians have a financial relationship (e.g., an ownership interest or a compensation arrangement) for the provision of "designated health services"²⁵⁹ except in a number of specifically exempt circumstances (e.g. where the physician is an employee of the entity). In addition to the federal law, a number of states have enacted laws or regulations which also prohibit some types of self-referrals, including services reimbursable by private health plans.²⁶⁰ The Stark law and similar state self-referral statutes or regulations are intended to avoid having financial considerations influence physicians' referral decisions. However, under a system that bundles payments to physicians and hospitals (or to physicians and other types of entities) to enable and encourage the delivery of coordinated services, physicians will inherently need to refer their patients to the other providers with which they have the bundled payment arrangement, and this may violate state and/or federal self-referral laws or regulations. Moreover, because the laws or regulations typically have exemptions for employment arrangements, they can create a disadvantage for organizational structures in which physicians are independent compared to health systems that employ physicians.
- **Prohibitions against payments in return for referrals of patients.** The federal Anti-Kickback statute²⁶¹ makes it a felony for any person to knowingly and willingly offer, solicit, or receive any remuneration for either referring a patient for an item or service, or for arranging or recommending an item or service, paid in whole or in part under a federal health care program. Many states have also enacted anti-kickback statutes or regulations.²⁶² The federal Anti-Kickback statute and state anti-kickback laws can make it illegal to create a program to reward physicians for following specific guidelines or to share the savings from the use of particular drugs or devices that have lower costs and higher quality.
- **Prohibitions against payments to physicians to reduce or limit services.** The Civil Monetary Penalty statute²⁶³ imposes financial penalties on hospitals that make payments to physicians as an inducement to reduce or limit services to Medicare or Medicaid beneficiaries. The law has been interpreted by the Office of Inspector General (OIG) at the Department of Health and Human Services (HHS) as prohibiting

such payments even if the services being reduced are not medically necessary or appropriate.²⁶⁴ Consequently, gain-sharing programs designed to reward physicians for reducing unnecessary services or unnecessary elements of services could make a hospital liable for civil money penalties. Although the law applies only to Medicare or Medicaid beneficiaries, the OIG has interpreted the law as also prohibiting such payments even for commercially insured patients, since the assumption is that incenting changes in practice for commercial patients would likely also result in changes in practice for Medicare or Medicaid patients, or that the amounts of payment incentives for changing practices, even though applied only to commercial payments, are set at levels designed to incent the changes for all patients.²⁶⁵

ix. Unwillingness to Make Reductions in Costs or Improvements in Quality

Finally, some providers may simply be unwilling to participate in the APM because they are succeeding financially under the current fee-for-service system and they project having lower profits under the APM.

b. How to Encourage Provider Participation in APMs

An APM cannot succeed unless providers participate, and APMs are far more likely to be successful if providers are participating willingly. Rather than trying to force providers into APMs they find problematic, it makes sense to design APMs in ways that avoid the problems described above. Approaches for doing this include:

- **Involve providers in the design of APMs.** It is not surprising that Alternative Payment Models designed solely by payers are more likely to be viewed favorably by payers than by providers. Many of the characteristics of APMs that are problematic for providers are viewed as desirable by payers, either because they simplify administration for the payer or because they shift more of the payer's financial risk to the provider. In order to attract participation by providers, an APM must be designed in a way that enables a provider to deliver high-quality care to patients in a way that is financially feasible for the provider. In order to attract participation of both payers and providers, and in order to successfully address both spending and quality, compromises will be required in the APM design between what would be ideal from the payer's perspective and what would be ideal from the perspective of the providers who will deliver and be paid for services through the APM.

It is far more likely that successful compromises will be found if APMs are designed through a collaborative approach involving both payers and the providers who would actually be implementing the APMs. Inviting providers to serve on an advisory committee is far less likely to be successful than creating a decision-making body that includes providers as well as the payers. Moreover, the providers who are involved in the design and decision-making need to include those who would actually be paid through the APM. In particular, APMs are more likely to be feasible for and

attractive to small providers if they are designed with involvement of small providers. However, special efforts have to be made to involve small providers because they cannot take time away from delivering patient services as easily as representatives of large provider organizations.

- **Design APMs using Regional Health Improvement Collaboratives or with state government oversight.** If multiple payers will be implementing similar APMs, it will be impractical to have each payer involving the same providers in separate design processes. However, if different providers are involved in separate APM design efforts, the result is likely to be unnecessary differences in the APM designs that will make participation unnecessarily complicated for providers who have patients insured by different payers. In a number of communities, non-profit multi-stakeholder Regional Health Improvement Collaboratives have brought multiple payers and providers together to agree on a common payment methodology to support innovative care delivery programs or on a common set of quality measures that all of the payers will use.²⁶⁶ Although concerns are often raised about whether such discussions violate antitrust law, antitrust prohibitions focus primarily on payers agreeing on a common *price* for services, not on developing a common *method* of payment.²⁶⁷ An alternative approach is for states to supervise the development of a common payment methodology, using the state action exemption under federal antitrust law.
- **Use the same definitions and measures in APMs that are intended to do the same thing, but don't force standardization where it is inappropriate.** There is no one "right" way to develop any APM, so it is not surprising that different payers might develop different APMs to try and achieve the same goals. However, no provider wants to deliver care differently to patients simply because they have a different payer, and administrative costs are higher if coding and billing are different for different payers. In most cases, it is unlikely that a small difference in one APM design compared to another will result in a major difference in spending or quality, so it makes sense to avoid creating unnecessary differences in APMs.

However, this does not mean that APMs should be designed in the same way when they are being used to achieve different goals or are being used with patients who have very different needs. For example, evidence-based standards for appropriate care can differ for patients of different ages, and outcome goals will likely differ for younger patients and older patients, so payers who insure different types of patients may need to use different APMs or to use different measures of quality. Differences in APM design may also be needed in different communities because of differences in the types of patients cared for by the participating providers and/or differences in the services available in the communities where the patients and payers are located (e.g., the more limited array of services available in rural communities.) If the quality measures or payment amounts in an APM do not match the specific patient needs and resources available in the community, the APM may not

be successful. Designing APMs through a collaborative process involving both providers and payers will help each understand the impacts on each other of differences in requirements so that reasonable compromises can be reached.

- **Enable providers to access claims data or other sources of information on the services their patients are receiving that are relevant to the APM.** Healthcare providers will be better able to determine whether to participate in an APM if they have the data needed to evaluate its impact and if they have confidence that they will be able to obtain the necessary data for accountability measures in a timely and accurate fashion after they begin participating. Data from only one payer will generally not be sufficient to accurately measure performance unless the payer represents a large portion of the provider’s patients.
- **Encourage other payers to participate in Medicare APMs both before and after providers begin participation.** Rather than only choosing regions where payers agree to participate, CMS could include any practice that is willing to participate in an APM just with Medicare participation, while also making individuals and purchasers with other types of health insurance aware of the benefits of the APM so they would encourage other payers to also implement the APM. Purchasers and smaller payers will be more likely to support implementation of an APM if a large payer such as Medicare has done so and if the providers in the community have already demonstrated a willingness to implement the APM.
- **Enable Medicare to participate in APMs that are being used by private payers.** Some private payers and providers have developed innovative APMs, but it may be difficult for the providers to participate or to succeed unless Medicare is also paying in a similar way. CMS could create a process for Medicare to participate in APMs that are being implemented by other payers if

there is a business case analysis showing there are opportunities for savings and care improvements for Medicare beneficiaries similar to those for other patients.

- **Reduce the inherently higher financial risks for providers during the initial implementation period for the APM.** When the APM is first implemented, providers will likely need to incur a variety of one-time costs to recruit and train staff or purchase new equipment, and they will likely be less productive in delivering new approaches to patient care and less successful in improving outcomes until their processes for care delivery have been refined. The combination of higher costs, lower revenue, and weaker outcomes during this initial phase can result in significant financial losses for a provider under an APM, even if the care delivery model will ultimately be financially viable. These initial risks can be reduced through actions such as:
 - ◆ Paying for a portion of the startup costs incurred by providers during the initial implementation of APMs.
 - ◆ Using less aggressive spending or quality targets initially.

Example: In the Advance Payment ACO Model, CMS provided 36 small physician-based Accountable Care Organizations participating in the Medicare Shared Savings Program with a total of \$67.8 million in up-front payments to invest in resources needed to improve care delivery. Each ACO received a fixed payment of \$250,000 plus an up-front payment of \$36 times the number of its historically assigned beneficiaries, and the ACO then received an \$8 per beneficiary per month (PBPM) payment for 24 months. The payments were recouped from any shared savings payments earned by the ACOs, but were treated as grants otherwise.²⁶⁸

Reasons Why Providers Don't Participate in APMs	Ways to Encourage Provider Participation in APMs
<ul style="list-style-type: none"> • Payers incur administrative costs to Implement APMs that weaken the business case for the APM • Reducing healthcare spending reduces profits for health insurance companies • Payers benefit by being a “free rider” while other payers implement an APM • Contracts with providers prohibit implementation of APMs 	<p>APM Design</p> <ul style="list-style-type: none"> • Design APMs to work within existing payer administrative systems • Use a common approach to coding and definitions in APMs • Design APMs that work for self-insured purchasers <p>Government Actions</p> <ul style="list-style-type: none"> • Require payers to publicly disclose the payment methods they use • Prohibit provisions of payer-provider contracts that limit the ability to implement APMs <p>Purchaser Actions</p> <ul style="list-style-type: none"> • Select insurance plans based on willingness to implement APMs • Contract for insurance and care delivery through purchaser coalitions • Use direct purchaser-provider contracting <p>Payer Actions</p> <ul style="list-style-type: none"> • Refuse to contract with payers that do not implement APMs • Develop the capability to contract directly with purchasers or to sell insurance products

- **Revise laws and regulations that create barriers to implementing APMs.** If there are laws and regulations that inappropriately create barriers to implementing an APM, changes in those laws or regulations may be needed. If a federal or state agency has the authority to waive the requirements in specific circumstances, that could be used as a temporary solution, particularly during the initial phases of testing and implementation of an APM. The statutes creating the Center for Medicare and Medicaid Innovation, the Medicare Shared Savings Program, and the Health Care Quality Demonstration Program each give the Secretary of Health and Human Services the ability to waive certain other requirements of federal law if necessary for the purposes of an APM.
- **Refuse to use providers that do not participate in the APM.** If a provider refuses to participate in the APM because it is unwilling to change care delivery or reduce excess profits, payers, purchasers, and patients could simply refuse to use that provider's services. If the provider is the only source of a particular service in a community, the payers, purchasers, and patients in that community may need to take steps to create alternative sources of receiving the service, such as:
 - ◆ Paying for patients to travel to other communities to receive services;
 - ◆ Providing financial assistance to assist new providers to begin delivering the service in the community.

3. Encouraging Participation by Patients

Willingness by payers and providers to implement an APM is necessary but not sufficient to ensure successful implementation. If patients do not have insurance from payers who are participating in the APM, or if they don't use the physicians and other providers who are participating in the APM to address their healthcare needs, then the APM will have a smaller impact on spending and quality than it would otherwise.

a. Why Patients May Be Unwilling to Participate in APMs

The fact that an APM is viewed favorably by payers or providers does not necessarily mean it is desirable from the perspective of the patients who would be receiving healthcare services supported by the APM. To date, most APMs have been designed primarily to reduce healthcare spending for payers, and the costs that patients incur and the quality of care they receive is at best a secondary consideration. Efforts to reduce the administrative burdens of APMs to make them more attractive to providers have often resulted in even weaker efforts to maintain or improve the quality of care for patients.

For example, a patient will be understandably concerned about an APM that:

- forces the patient to receive their care from a narrow list of providers that were selected based primarily on the price the providers were willing to charge rather than the quality of care they committed to provide;
- requires the patient to pay more in cost-sharing than they would have paid under the fee-for-service system for the specific services they receive;

- financially penalizes the patient's physician if the physician has to order more services or more expensive services to meet the patient's needs;
- financially rewards a provider if that provider delivers fewer services than the patient needs;
- requires the patient to pay for services even if the quality of care the patient received is poor, because the quality of care for other patients was acceptable;
- fails to evaluate the outcomes achieved or the quality of care delivered for the specific types of health problems the patient has.

Many proposed "population-based payment" systems have many or all of these undesirable characteristics.²⁶⁹ Patients have had the choice of participating in these types of payment models for many years through HMOs and capitated physician groups, and participation in those models has decreased rather than increased.

Many "bundled payment" and "episode payment" models also have the same characteristics. For example, in the Medicare Comprehensive Care for Joint Replacement Program, a patient has the ability to choose which hospital and surgeon will perform the surgery, and the patient nominally retains the ability to choose which home health agency or skilled nursing facility will provide post-acute care services. However, the target prices set by CMS for the entire joint replacement episode are not adjusted if a patient has characteristics that would make the patient more likely to need a skilled nursing facility, so the hospital will be financially penalized if it orders a SNF stay rather than sending the patient home with only home health agency support regardless of whether the patient needs a SNF stay or not.²⁷⁰

A patient's unwillingness to participate in these kinds of APMs does not mean that the patient feels the current fee-for-service payment is ideal. A patient may simply prefer to retain the ability to choose the highest-quality providers even if it results in less coordinated care, and they may prefer to pay more to receive care from providers who have the flexibility to order the services the patient needs rather than using providers who charge less but do not take accountability for the patient's outcomes.

b. Why Patients May be Unable to Participate in APMs

At the other extreme, some patients who could potentially benefit from an APM might be *unable* to do so if the design of the APM would cause providers to lose money caring for those patients. A provider participating in an APM will be understandably concerned about accepting a new patient who has multiple, unusual, or complex needs unless the APM:

- provides higher payments to the provider to cover the costs of the additional time or resources needed to care for that patient;
- excludes or adjusts for the legitimately higher utilization or spending on the patient when determining penalties or bonuses for utilization/spending;
- excludes or adjusts for any differences in care delivery or outcomes that are appropriate or feasible for the patient when determining penalties or bonuses based on quality.

Most current value-based payment systems and APMs do not have these characteristics. For example, most APMs do not risk-adjust payment amounts, spending measures, or quality measures based on characteristics of the patient such as functional status, availability of caregivers, language skills, and access to food, shelter, or transportation that can significantly affect the types of services patients need and the outcomes they can achieve with any particular level of healthcare services. Most APMs do not exclude or provide additional payments for “outlier” patients who have unique characteristics or circumstances that cause unusually high levels or utilization/spending or where standardized measures of care quality would be inappropriate.²⁷¹ For example, the CMS Comprehensive Care for Joint Replacement program does not adjust payments for patient who have characteristics that make them more likely to need expensive forms of post-acute care. This creates a financial incentive for hospitals and surgeons participating in the program to avoid higher-need patients.

A growing number of studies have shown the kinds of negative impacts that these payment models can have on providers who treat a disproportionately high number of such patients.²⁷² Studies have shown that Medicare Advantage plans have methods of “cherry-picking” patients and “lemon-dropping” patients based on characteristics that can affect spending or quality but that are not adjusted for in the population-based payment system used to pay those plans²⁷³, and similar problems could arise when similar population-based payment methods are used to pay providers rather than insurance plans. This could increase rather than reduce disparities in care.

c. Encouraging Patient Participation in APMs

If APMs are going to be attractive to patients who have choices, they need to be designed to benefit the patients, not just payers and providers. If patients with higher needs are going to be attractive to providers in APMs, the APMs need to be designed so as to not penalize the provider for taking care of those patients.

These are not separate problems, but rather two sides of the same coin. The solution to both problems is to design an APM to be as *patient-centered* as possible. A patient-centered APM design that addresses both problems would have the following characteristics:

- **Payment amounts based on patient needs:** if the amounts paid to a provider under the APM are based as much as possible on the patient’s needs, then providers can treat patients appropriately without being financially penalized for doing so, and patients can feel comfortable they will receive the care they need from the providers participating in the APM.
- **Accountability focused on avoidable utilization and spending:** if accountability is tied to specific aspects of care where it is known that savings can be achieved in ways that benefit, or at least do not harm, patients, then providers will not be rewarded for withholding needed care, and patients can feel comfortable that the provider’s decisions are based on the patients’ needs, not on financial rewards.
- **Accountability focused on the quality of services and outcomes for the individual patient:** if providers are only paid for the care to an individual patient when that individual patient receives care that meets standards specific to that patient’s characteristics and needs, each patient can feel comfortable that the care they receive will be of high quality, regardless of what happens to other patients.

In many cases, individual patients are the actual purchaser of the services they need, not a third-party payer or other purchaser. This includes people without insurance but also the growing number of individuals with high-deductible health insurance plans who find that they are paying for 100% of the costs for all or most of the services they receive. Patient-centered payment allows the use of direct patient-provider contracting, since the payment to the provider is based on the specific patient’s needs and the provider’s accountability is based on costs and outcomes for that specific patient.

Reasons Why Patients Don't Participate in APMs	Ways to Encourage Patient Participation in APMs
<p>Unwillingness to Participate</p> <ul style="list-style-type: none"> • Use of narrow networks based on prices rather than quality • Higher cost-sharing than under fee-for-service • Penalties for ordering services patients need • Rewards for withholding needed services • Payment required even if the quality of care the patient receives is poor • Failure to evaluate quality or outcomes for the patient’s specific health problems <p>Inability to Participate</p> <ul style="list-style-type: none"> • Inadequate payments for patients who need more time or resources • Failure to adjust for necessary utilization when determining bonuses/penalties for spending • Failure to adjust for feasible outcomes when determining bonuses/penalties for outcomes 	<p>APM Design</p> <ul style="list-style-type: none"> • Adjust payment amounts based on differences in patient needs • Base accountability on measures focused on avoidable utilization and spending • Base accountability on measures focused on patient-specific quality and outcomes

B. Finalizing the APM Parameters

1. Challenges in Finalizing the Initial Parameters for the APM

Most of the structural elements of the APM and methods of operationalizing them can be determined prior to implementation using the approaches described in Sections VI and VII. However, in many cases, it will be difficult to specify the “right” payment amounts and targets for spending and utilization before the APM is actually implemented.

One reason for this is that in many cases, the payment amounts and targets will depend on which specific payers, providers, and patients are participating, and this will not be known until after they agree to participate. For example, the CMS BPCI APMs have been implemented in two phases – in the first phase, a provider organization makes a preliminary commitment to participate, and historical data on utilization and spending are then generated for that provider’s patients so that both CMS and the provider can determine the actual parameters of the APM. Then, if the provider commits to receive the payments, the payments are actually made in a second phase.²⁷⁴

However, a second reason why it is difficult to fully specify the parameters of the APM before it is implemented is that this requires information that can only be obtained from providers that are delivering services in a different way, but providers cannot deliver services in that way without having an alternative payment model to support them. For example:

- **Determining appropriate payment amounts for new or different services.** If an APM is intended to support the delivery of a service that is not currently eligible for payment under current payment systems, the APM will need to specify how much will be paid for that service. However, it is difficult to estimate the cost of such a service if there is little or no experience in delivering the service due to lack of payment. For example, a payment model might be designed to pay a non-clinician educator to educate a chronic disease patient about how to avoid exacerbations, but it will not be clear how many patients can be adequately educated by a single individual, how much will need to be paid for an educator with the skills necessary to be effective, etc. until the APM is actually implemented.
- **Setting payment amounts for bundled services.** If an APM provides a bundled payment that replaces one or more current fee-for-service payments and also provides flexibility to deliver services that are not currently eligible for payment, the APM needs to specify how much will be paid for the bundle. However, it is difficult to estimate the appropriate payment amount without an understanding of how often current services would be replaced by new services, the extent to which fixed costs supporting existing services can be eliminated,

etc. For example, many physicians would prefer an APM that replaces current Evaluation & Management payments (which are mostly limited to face-to-face visits with a physician) with a monthly payment that would provide the flexibility to schedule patient phone calls with the physician instead of just office visits, to make contacts with patients using nurses instead of physicians, etc. However, it will not be clear what amount these monthly payments should be until it is determined what proportion of office visits can be eliminated, what types of additional staff the physician practice would need to hire, etc., and those changes cannot be made until the APM is actually implemented.

- **Defining methodologies for risk-adjusting/stratifying payments.** An APM that creates a bundled payment in place of fees for individual services will likely require a method of stratifying or adjusting the bundled payment amount to reflect differences in patient needs. However, the patient characteristics that affect the level of services may not be adequately captured by ICD-10 diagnosis codes. The APM would need to specify what combination of patient characteristics would be associated with each payment stratum and how much the payment amount for those patients would be, but it is difficult to do either of these things without data on how many patients have particular combinations of characteristics and how the appropriate services will differ for different characteristics. For example, an APM might create a monthly payment to support home-based palliative care services to patients, but the payment amounts would need to be higher for patients with lower functional status, less caregiver support, etc., and it will not be clear how many patients have those characteristics and how many patients in each category could be managed by a palliative care team until the APM is actually implemented and new codes are available for recording this information.
- **Setting standards for performance on outcomes and other quality measures.** As discussed in Section VI, although there is broad agreement that it would be desirable to have APMs that are designed to improve patient outcomes, there is little outcome data available that can be used for establishing baseline levels of outcomes and performance standards because of the significant costs involved in collecting outcome data and the lack of a business case for providers to incur those costs under current payment systems. For example, if an APM that pays for managing knee or hip osteoarthritis holds providers accountable for addressing pain and mobility problems, data on expected levels of pain and mobility would not be available until they were collected through implementation of the APM.

The more innovative the APM – i.e., the more than it differs from the current payment system – the more likely there will be a need for initial beta testing and potentially for additional rounds of refinement after the APM is implemented more widely.

These problems create a “chicken and egg” conundrum for implementation; payers and providers will be reluctant to participate in the APM without knowing for sure what the parameters will be, but the data needed to determine the parameters will not be available until after the APM has been implemented.

2. Using a Beta Testing Process for APMs

In other industries, new products do not immediately jump from the design table to full-scale production. One or more prototypes are created and tested on a limited scale to identify problems and opportunities for improvement. The design is then revised before broader production begins. In addition, product designs continue to be refined and new versions of products are created whenever the benefits for consumers are expected to outweigh the costs of making the changes.

A similar “beta testing” process is needed for many APMs. The more innovative the APM – i.e., the more than it differs from the current payment system – the more likely there will be a need for initial beta testing and potentially for additional rounds of refinement after the APM is implemented more widely. To be successful, beta testing will likely require:

- **Participation by a limited number of interested providers.** Initial participation by a small number of providers could allow sufficient data on outcomes, costs, and patient characteristics to be gathered and analyzed in order to refine the key parameters of the APM. However, providers that participate would need to be adequately diverse to ensure that the data would be representative of the providers that would participate when the APM is implemented more broadly.
- **Use of “best estimate” parameters to initiate APM testing.** In order to implement the APM in the beta test phase, initial values for the payment amounts and accountability targets will be needed even though no one can yet be sure whether they are “right.” There will generally be some information available to help with this, such as the information used to outline the business case in Section IV, but in some cases, educated guesses may be needed until initial data are available.
- **Protection for providers, payers, and patients against financial harms.** Although it is likely that the initial parameters will be at least somewhat “wrong” (and potentially very wrong), it will not be clear in which *direction* they are wrong, e.g., whether the payment amounts will be too high or too low, and whether the performance standards will be too hard or too easy to achieve. The goal should be to achieve some benefits while avoiding significant negative impacts for all stakeholders – patients, providers, and purchasers – during the beta testing period. This will require a monitoring and reconciliation process to ensure that no one is financially harmed while the parameters are still being refined.
- **Resources for data collection by providers.** Because the providers in the beta test will be expected to collect data needed to improve the design of the APM,

they will need to receive adequate resources to collect those data.

It is important to recognize that:

- the goal of the beta testing process will be to revise the *design of the payment model*, not merely to deliver services in a better way. If the goal were merely to determine the best way to implement a particular approach to care delivery, providing grant funds would be an easier and more effective approach, since modifications to the care delivery approach would at most require adjustments to the budget allocations of the grant. However, since the goal is to implement a *better payment system*, there needs to be a capability for adjusting the way the providers are paid if it becomes clear that there are problems with the initial design.
- the purpose of beta testing is to *refine* the APM, not to evaluate whether it “works.” In fact, it is likely that an evaluation conducted before an APM has been adequately refined will conclude that the APM is less effective in reducing costs or improving quality than it would ultimately be, which could cause it to be terminated prematurely or discourage other payers or providers from implementing it.

C. Evaluating the APM

Even with the most careful planning and design, an Alternative Payment Model may not be as successful as was hoped, or it may create unanticipated problems. As with any new product or service, it makes sense to carry out an evaluation of any newly implemented APM in order to identify any problems and make revisions in the APM to correct the problems.

Federal law requires that APMs created by the Center for Medicare and Medicaid Innovation be evaluated, and the law further requires that if the evaluation shows that the APM is not achieving savings while maintaining quality or is not improving quality without increasing spending, the APM must be modified or terminated. The law also allows CMS to expand the use of successful APMs to more providers and patients if the Office of the Actuary certifies that the APM will not increase spending.

In the private sector, many purchasers and payers have indicated that they are unwilling to implement, continue, or expand the use of an APM without either an evaluation of the APM or an actuarial certification indicating that use of the APM will reduce spending.

Consequently, determining whether to implement, continue, and expand APMs requires both structuring evaluations correctly and interpreting their results properly.

1. Defining What Should Be Evaluated

Since the motivation for creating APMs has been to reduce healthcare spending and improve quality and outcomes for patients, it is natural to assume that an evaluation should focus solely on determining whether spending was lower and/or quality and outcomes were better after the APM was implemented. However, in general, an Alternative Payment Model does not *directly* reduce healthcare spending or improve the quality of care. The amount spent on healthcare services, and the quality and outcomes of those services, are determined by the types of needs that patients have and the way that healthcare providers address those needs. An APM may be a *necessary* condition for achieving reductions in spending or improvements in quality, but it will rarely be *sufficient*. Implementation of an APM will result in lower spending and equal or better quality if, and only if, all six of the following conditions are in place:

- (1) there are opportunities to reduce spending without harming patients and/or opportunities to improve quality without increasing spending,
- (2) physicians and other healthcare providers know how to change care delivery in order to achieve those opportunities;
- (3) the current payment system creates barriers to making those changes in care delivery;
- (4) the alternative payment model is able to remove the barriers;
- (5) the providers actually make the changes in care delivery once the barriers are removed; and
- (6) the changes in care delivery have the expected effect on spending and quality.

If the APM fails to successfully change the aspects of payment that were viewed as barriers to success, then failure to achieve savings or improve quality can be appropriately ascribed to the APM. It is important to evaluate whether the payment changes have actually enabled changes in care delivery, rather than simply determining whether there have been changes in spending and quality, because savings or changes in quality could be due to changes in care delivery or other factors unrelated to those that the APM was intended to support.

On the other hand, if the APM successfully removes the payment barriers it was intended to remove, but one or more of the other conditions isn't present, then it isn't appropriate to blame the *payment model* for failure to achieve savings or improve quality. Rather than modifying or terminating the APM, it may make more sense to determine whether additional actions besides the change in payment are needed to support the desired outcomes. For example, if physicians or other providers do not make changes in care delivery even after the payment barriers are removed, there may be barriers other than the payment system that need to be addressed, such as fears of malpractice liability if fewer tests are ordered, or inability to find employees with the necessary skills. An alternative to terminating the APM altogether would be to narrow the use of the APM to situations in which the favorable conditions do exist.

In some cases, there will already be extensive evidence regarding the feasibility and effectiveness of the change in care delivery the APM is intended to support. For example, grant-funded demonstration projects in which those changes were made may have been evaluated and shown to be effective.²⁷⁵ However, in other cases, it may have been impractical to test the impact of the changes in care delivery without an appropriate payment model to support those changes. In these latter situations, the evaluation will not only need to determine whether savings or improvements in quality have occurred, but also to determine the causes of the changes or lack of changes. This requires two separate components to the evaluation – one to determine whether the APM removed the payment barriers to changing care delivery, and another to determine whether the change in care had the intended impacts on cost, quality, and outcomes.

It is also important to recognize that the amount of savings and the level of improvement in quality that can be achieved depends heavily on the baseline level of spending and quality for the specific providers and patients who are actually participating in the APM. In many cases, the providers who are most likely to be the early adopters will be those who have already been able to implement the desired changes in care through special, short-term arrangements, such as demonstration grants. For them, the APM is needed not so that care can be *changed*, but so that changes that have already been made can be *sustained*. In these cases, it is particularly appropriate to focus the evaluation on whether the APM successfully removes the barriers in the current payment system, since an evaluation of the impact of the care change would presumably already have

been completed or be underway as part of the demonstration project. Moreover, savings may have already been achieved through the grant funding, and so it will be important to establish the Target appropriately.

2. Defining the Timeframe for Evaluation

It would be ideal if at least a preliminary evaluation of a new APM could be completed quickly, since that would allow any needed changes to the APM to be made quickly, particularly if the APM is causing spending to increase or is harming the quality of care for patients. However, it seems unlikely that early evaluations – e.g., evaluations completed within the first year or even the first two years following implementation of the APM – will accurately describe the longer-run impacts of an APM that makes major changes in the way providers are paid and/or an APM that is intended to support large changes in the way care is delivered. This is because it will take more time to make the changes as planned and there will likely be more unexpected issues arise that need to be resolved.

On the other hand, it is problematic to continue using an ineffective APM after successive evaluations have shown negative results, based only on the hope that positive results will be achieved with “a little more time.” The longer an ineffective APM is allowed to continue, the longer it will take to develop something different and then implement it.

Instead of selecting arbitrary periods of time over which to pilot test and evaluate APMs, it would make more sense to base the evaluation timeframe on an analysis of the amount of time that will be needed to implement the care delivery and payment changes. That analysis can also be used to establish clear deadlines and milestones for progress that payers and providers can be held accountable for achieving.

3. Setting Thresholds for Success

a. Using Thresholds Based on Statistical Significance

Evaluations of Alternative Payment Models are typically designed to assess whether there was a statistically significant difference in the changes in spending and quality/outcomes of care for the patients who are participating in the APM and for a comparison group of patients who were not participating in the APM.

APMs Intended to Reduce Spending

If the goal of the APM is to reduce spending, success is usually defined as a reduction in spending on patients participating in the APM that is larger, by a statistically significant amount, than any reduction in spending for non-participating patients.²⁷⁶ However, defining success this way is problematic for two reasons:

- If the APM is being tested with a relatively small group of providers and patients, it will be difficult to achieve statistical significance on a traditional frequentist significance test even when relatively large amounts of savings have occurred, solely because of the small sample size and the normal random variation in spending for the patients. The higher the confidence level sought by the evaluator, the more likely it is that the APM will be declared a failure even when success has actually occurred, i.e., the lower the Type I error rate that is desired, the higher the Type II error rate.
- Conversely, the larger the number of providers and patients participating, the easier it will be for an evaluation to find savings that are “statistically significant” even though the dollar amount of savings is very small.

As explained in more detail in Section VI, the broader the measure of spending that is being evaluated, the greater the variability in the patients participating in the APM, and the more difficult it will be to find a comparison group. This makes it less likely that an evaluation will conclude that savings are statistically significant.

As a result, defining success in terms of statistical significance creates the risk that a decision to continue or terminate an APM will be determined more by the number of patients participating in the initial test of the APM than on the actual amount of savings the APMs could ultimately produce. As discussed in the previous section, APMs that make more dramatic changes in

payments will likely need to be implemented on a more limited scale initially in order to refine the parameters of the APM. Even if savings are produced during the limited-scale testing, it is unlikely that the savings will reach typical levels of statistical significance.

APMs That Are Intended to Improve Quality

A focus on statistical significance is even more problematic if the goal of the APM is to improve quality and outcomes without increasing spending. Proving that an APM did not increase spending is difficult because even in the absence of the APM, spending could be higher in any given year solely due to random variation.

As a practical matter, it is unlikely that an actuary would ever certify that an APM “did not increase spending” unless the APM actually *reduced* spending by enough on average to offset most of the increases in spending that can occur due to random factors. However, this is unlikely to happen if the APM is not designed to achieve savings. Moreover, the smaller the number of patients involved, the more likely it is that random increases in spending will be large, so the less likely it is that the APM will be able to “prove” that it didn’t increase spending.

The same issues arise in evaluating changes in quality. If the APM is supposed to improve quality, actual im-

Improvements in quality may fail to reach statistical significance when there is variation in quality across patients and/or a small number of patients are participating in the APM.

b. Defining Success in Evaluations Differently

There is growing recognition that the “cult of statistical significance” in evaluation research can easily lead to bad decisions.²⁷⁷ Similarly, traditional actuarial evaluations of payment models tend to favor the status quo²⁷⁸, even though it is clear that the status quo needs to be changed.

Rather than focusing solely or primarily on statistical significance, an evaluation of an APM’s impact on spending should assess three factors:

- the magnitude of the actual savings achieved;
- the probability that the true savings are greater than zero;
- the probability that the true savings are greater than a minimum desirable level (or that an increase in spending was less than a maximum desirable level).

Similarly, an APM’s impact on quality should be assessed on three factors:

- the change in quality or outcomes that was actually achieved;
- the probability that the true change in quality/outcomes was different than zero; and
- the probability that the true change in quality/outcomes was greater than a minimum desirable level (or that a decrease in quality/outcomes was less than a maximum desirable level).

Success should then be determined by the balance among those factors.

Example: In the Medicare Shared Savings Program, an ACO with 5,000 assigned patients must achieve a 3.9% reduction in spending in order to be determined to have achieved “true” savings, whereas an ACO with 50,000 assigned patients needs to only achieve a 2.2% reduction in spending. This means that if there are ten ACOs, each with 5,000 patients, and they start with average Medicare spending of \$12,000 per patient and reduce that spending by 3.5%, the resulting \$21 million in savings ($3.5\% \times \$12,000 \times 5,000 \times 10$) would be considered “non-significant.” However, if an ACO with 50,000 patients started with \$9,000 in spending per patient and spending decreased by 2.2%, the \$9.9 million in savings ($2.2\% \times \$9,000 \times 50,000$) would be considered “significant” even though it is only half as much. Instead, assessing the ACOs based on both the size and significance of their savings would allow the smaller ACOs to be recognized for having achieved greater savings.

4. Taking Actions Based on the Evaluation Results

Several different actions could be taken based on the results of an evaluation:

- Continuing or expanding the use of an APM;
- Modifying the APM;
- Terminating the APM; or
- Continuing the payment model through a different mechanism.

a. Expanding the Use of an APM

Federal law permits the Secretary of Health and Human Services to “expand the duration and scope” of an APM implemented by CMMI, including “implementation on a nationwide basis” if:

- the Secretary determines that the expansion is expected to reduce spending without reducing the quality of care or improve the quality of patient care without increasing spending; and
- the CMS Chief Actuary certifies that the expansion would reduce (or would not result in any increase in) net Medicare spending.²⁷⁹

Consequently, if the evaluation indicates that the spending is lower and quality has been maintained or improved under the APM, and that those results would be expected to continue, CMS could continue implementing the APM as it is currently structured with the current providers. The APM could also be expanded to a larger number of participants, including potentially all eligible patients and providers.

However, an evaluation showing that the initial participants in the APM have reduced spending without harming quality (or have improved quality without increasing spending) does not necessarily mean that additional participants would achieve similar results. Moreover, it is not necessarily clear whether additional participants would do better or worse than the initial participants. There are several reasons for this:

- The initial participants may have been those who had the biggest opportunities to achieve savings and/or improve quality. In this case, future participants may produce smaller results even with similar effort, simply because they have less opportunity to reduce spending or improve quality;
- The initial participants may have already achieved savings or improved quality through grant funding or simply on their own initiative, but they required a change in payment in order to sustain those achievements. In this case, future participants might achieve even more than those who participated initially.
- The initial participants may have had an unusually high level of commitment and passion to make changes in care delivery, or special resources at their disposal. In this case, subsequent participants might devote less time or resources to making the APM a success, and the results would be worse.

Problems with Mandatory Tests of APMs

Concerns that the providers who will voluntarily participate in an APM are not representative of other providers have led to calls for APMs to be implemented on a mandatory basis. In the Medicare program, this means that CMS would require providers to participate in an APM whether they wish to or not. CMS is using this approach in the Comprehensive Care for Joint Replacement APM, although on a more limited basis than originally intended. It announced plans to implement other APMs with mandatory participation but later withdrew them.

“Mandatory” demonstrations are undesirable for several reasons:

- In an effort to mimic a randomized control trial, most proposals for mandatory demonstrations require that a subset of providers and patients participate while others would be prohibited from doing so. If the potential benefits of the APM are sufficiently great to justify requiring every patient in a region to be part of it, at least some patients in other parts of the country would be penalized if providers in their region are precluded from participating.
- On the other hand, there is no guarantee that the APM will be successful, otherwise there would be no need for an evaluation. Moreover, there may be good reasons why specific providers are not willing to participate in a particular APM; APMs designed to reduce spending have the potential to harm patients and to harm certain kinds of providers.²⁸⁰ It is inappropriate to require providers and patients to participate in something they do not believe would help them or that they believe could potentially harm them before an evaluation has been completed. If it is believed to be desirable to randomize the APM test, that can be done without mandating participation; indeed, randomized control trials of drugs and medical devices recruit patients to participate, they do not force them to do so.
- In Medicare, because beneficiaries would retain freedom of choice about the providers they use and the services they receive, randomization could not be done at the patient level. The approach used in most mandatory demonstration proposals is to randomize by large geographic regions; while this may deter patients from traveling to a different geographic region to either participate in or avoid participating in care delivery under the APM, it would not preclude them from doing so. For example, there is anecdotal evidence that following initiation of the Medicare CJR

Successful APMs will not achieve savings simply because they pay in a different way or because they create an “incentive” to spend less; they will achieve savings by removing specific payment-related barriers to changing care delivery that will impact specific opportunities for improvement. Both the opportunities to achieve savings and the barriers to pursuing those opportunities will differ from community to community. This means that “one size fits all” APMs will be less likely to achieve the full amount of savings that are possible nationally or to provide better care to patients in all parts of the country than a more customized approach.

demonstration, higher-risk patients living in mandatory participation regions began to be referred to surgeons and hospitals located in non-participating regions.

- It is well-known that there are significant differences in the health status of patients and the practice patterns of providers across geographic regions. Because of the large size of the geographic areas that would need to be chosen in order to minimize inter-region referrals of patients, randomization would only have a limited ability to eliminate the influence of inter-regional differences on the results, so differences in savings and quality between the participating and non-participating regions would likely not be due solely to the APM.
- Because the mandate may only last until the evaluation is completed, providers who would be unwilling to participate voluntarily may also be unwilling to make the changes in care delivery necessary to achieve success. Consequently, the evaluation of a mandatory *demonstration* project would not necessarily show what the impacts of a mandatory *program* would be. Indeed, if a subset of providers did not

want the APM to be universally mandated, they would have an incentive *not* to produce successful results during the demonstration.

The implicit assumption that is made by those advocating for mandatory demonstrations is that if an APM is found to be successful, it should be mandated for use by every provider for every patient. However, nothing in the law requires this. The authorizing statute for CMMI says that the Secretary of HHS “may... expand (including implementation on a nationwide basis) the duration and the scope of a model that is being tested... to the extent determined appropriate by the Secretary.”²⁸¹ The law also does not clearly

define what “expanding the scope” of a model or “implementation on a nationwide basis” means.

In many cases, it will likely be both desirable and appropriate to create permanent but voluntary APMs. Successful APMs will not achieve savings simply because they pay in a different way or because they create an “incentive” to spend less; they will achieve savings by removing specific payment-related barriers to changing care delivery that will impact specific opportunities for improvement. Both the opportunities to achieve savings and the barriers to pursuing those opportunities will differ from community to community. This means that “one size fits all” APMs will be less likely to achieve the full amount of savings that are possible nationally or to provide better care to patients in all parts of the country than a more customized approach. Mandating a single

approach may lower administrative costs for Medicare and other payers, but it will also likely lower healthcare savings even more, thereby reducing the net benefits rather than improving them.

Congress has already created one permanent, voluntary APM – the Medicare Shared Savings Program. It is not a temporary demonstration project, but participation in it is voluntary for providers. Although it has had a very limited impact on Medicare spending, that is due to the problematic design of the payment model, not the voluntary nature of the program. A similar approach to voluntary participation could be used for other, better-designed APMs.

b. Modifying or Terminating the Use of an APM

An APM implemented by the Center for Medicare and Medicaid Innovation (CMMI) is not required to maintain or reducing spending initially. What the law requires is that the Secretary of HHS “terminate or modify the design and implementation” of an APM unless (i) the Secretary determines that the APM is expected to improve the quality of care without increasing spending, reduce spending without reducing quality, or improve quality and reduce spending, and (ii) the CMS Chief Actuary certifies that the model will maintain or reduce spending. The law specifies that termination “may occur at any time after such testing has begun and before completion of the testing.”²⁸²

While terminating an APM that has not shown the desired results might sound like a very prudent approach, it could have the perverse effect of reducing the likelihood of success. As noted in Section VIII.A, physicians, hospitals, and other healthcare providers are unlikely to fundamentally change the way they deliver care in response to a payment change that may only last a few years. Moreover, if it is known that a negative evaluation will result in termination of the APM rather than an effort to make improvements, the physicians, hospitals and other providers who are participating in the APM will likely be unwilling to make significant financial investments that could only be recouped over a long period of time or if there would be significant costs involved in reversing changes in care when the APM that supported the changes is terminated.

Medicare and other payers could mitigate this problem by making two commitments to providers:

- to continue implementing an APM for a long enough period of time to ensure that changes in care delivery can be fully implemented and to recoup the costs they incur in participating in the APM. The more complex the changes in care, and the more expensive the investments needed by providers, the longer the time period that an APM will need to stay in place.
- to modify the APM in an effort to correct any weaknesses before terminating it. As discussed in the previous section, it is essentially impossible to adequately define the parameters of a significantly different payment model until it has actually been implemented in at least a beta test. Multiple modifications may be needed before a determination can be made as to whether the APM can be successful.

c. Continuing the Payment Model Through a Different Mechanism

It is entirely possible that an evaluation will show that outcomes have improved but spending has also increased, and there may be no way to modify the APM to reduce spending in a way that would be viewed favorably by providers or patients would. Although this result would mean the payment model would not qualify for continuation as an APM, it does not mean that the payment model should be terminated. Instead, it could be continued through a different route.

For example:

- **Payments for New Services.** If a barrier that the APM was intended to address was lack of payment for specific services under current payment systems, payments for those services could be added to the existing payment system because of the improvements in outcomes the evaluation showed would result. For example, because of the recognized problems with paying only for face-to-face interactions between physicians and patients, and in response to evaluations of a variety of projects showing the benefits of paying for a broader array of services, CMS has changed the Medicare Physician Fee Schedule to pay for a series of additional services that do not require face-to-face interactions or interactions only with physicians. No APM was needed to do this.
- **Increased Payments for Existing Services.** If the barrier the APM was intended to address was inadequate payment for services under current fee schedules, the payment amounts could be increased based on the results of the evaluation.

As discussed in Section V.A, one key reason for using APMs to fill payment gaps rather than simply changing existing payment systems is that federal laws require that changes to the physician fee schedule and to hospital outpatient payments must be budget neutral, so an increase in payment for one service requires that payments for all other services be reduced. However, there are no similar rules requiring reductions in other payments when Medicare provides coverage for new medications or medical devices that will be more expensive than existing approaches. Consequently, if a change in services improves patient outcomes but cannot qualify as an APM, legislation may be needed to authorize higher spending to support that change.

5. Using Better, Faster Approaches to Evaluating APMs

The current approaches that are being used to evaluate APMs are contributing to the slow progress in implementing APMs. The process CMMI has used to carry out testing of APMs has been extremely slow, expensive for CMMI, and burdensome for providers. Requiring an elaborate evaluation be designed for every APM it implements has made it difficult to implement more than a few APMs and impossible to implement any of them quickly. This is one reason why most proposals for innovative alternative payment models that have been submitted to CMMI have not been implemented. In most cases, when a payment model is tested by CMMI, only a small number of providers are permitted to participate for many years even if there is broad interest in participation. Under most of the payment demonstrations that CMMI has implemented, 18 months or more have elapsed from the time an initiative is first announced to the time when providers actually begin to receive different payments. Many providers have decided not to even apply to participate in otherwise desirable CMMI programs and others have dropped out of the programs in the early phases solely or partly because of the cost and time burden of participating.²⁸³

In contrast to current value-based payment models, well-designed APMs would actually reduce the need for an evaluation to determine whether spending was lower and quality was higher. If the APM explicitly holds the

participant accountable for maintaining or reducing specific types of spending for a particular condition compared to a target and for maintaining or improving performance on quality standards and outcomes, much of what needs to be “evaluated” is built into the accountability components of the APM. If a provider is failing to maintain or reduce spending or failing to maintain or improve quality, the provider will either need to improve performance or stop participating in the APM. Rather than framing the decision in terms of continuing or terminating the entire payment model, the decision should be whether the payment model needs to be refined and whether individual providers should be permitted to continue participating.

Fortunately, Federal law permits this kind of approach to be used by CMMI. There are no limits in the law as to (1) how many providers can participate in testing, (2) how the evaluation should be conducted, (3) how quickly a determination is to be made as to whether the model improves quality or reduces spending, or (4) how often the design of a model can be modified before it is terminated or expanded. Under the law, the Secretary of Health and Human Services is authorized to modify the design and implementation of a model after testing has begun if the model is not expected to either improve quality without increasing spending or reduce spending without reducing quality.²⁸⁴

D. Revising the APM Parameters

1. The Need to Adjust APM Parameters Over Time

The initial parameters of the APM (i.e., the amounts paid for individual services or bundles of services, the utilization/spending targets, the quality targets, etc.) will all have been established based on the specific circumstances and information that existed when the APM was being designed and initially implemented. Over time, those circumstances and/or the information available will change. In particular:

- the costs of delivering services will change, due to increases in worker wages to keep up with inflation and changes in the prices of supplies, medical devices, drugs, etc. Although most costs generally increase over time, improvements in methods of manufacturing, introduction of competitive products, and more efficient approaches to delivering care may result in some services becoming less expensive over time;
- evidence about the causes and appropriate treatments of diseases may change. This may result in the need to use different approaches to diagnosis and treatment that may be more or less expensive, and there may also be changes in expectations about the outcomes that should be achieved for patients;
- new technologies for diagnosing or treating disease may be developed, which may increase or decrease costs and increase the expected level of outcomes that can be achieved;
- the size of a particular community or the availability of providers in the community may change, which can make it more or less difficult or expensive to obtain specific types of services;
- the prevalence or severity of health conditions may change, due to changes in lifestyles, the environment, preventive care, etc. Increases or decreases in the need for services can result in lower or higher average costs of delivering those services. Expectations about the types of outcomes that can be achieved may also need to change.

Failure to update the APM parameters appropriately in response to these changes could mean that the APM would no longer adequately enable and encourage the best quality care at the lowest possible cost. Moreover, if healthcare providers do not believe that appropriate adjustments will be made over time, they may be unwilling to participate in the APM at all.

The need for updates is not unique to APMs. Payment amounts in the fee-for-service system are typically revised annually, if not more frequently, by Medicare and other payers, although those revisions may or may not cover changes in costs. However, APMs also have mechanisms by which healthcare providers take accountability for spending and quality, and updates are needed to those mechanisms as well as to the amounts paid for the delivery of individual services or bundles of services.

The initial Spending and Quality Targets for an APM need to be set to ensure that either (a) spending under the APM will be lower than it would be otherwise, without harming quality, (b) quality will be better than it would otherwise, with no increase in spending, or (c) spending will be lower and quality will be better than otherwise. Until those Targets are achieved, they will need to be updated to reflect changes in spending or quality that would have occurred in the absence of the APM.

- For example, if spending would ordinarily have been expected to increase from the first to the second year of APM implementation due to general inflation in costs and prices, then the spending target in the second year of the APM would need to be higher than in the first year. Otherwise, the provider would be forced to achieve more savings in the second year than the first year (if the goal was to achieve savings) or to achieve savings for the first time (if the goal was to maintain spending while improving quality).
- If a new drug or procedure is developed that providers are using to improve outcomes for patients regardless of whether they are participating in the APM, then the quality and spending Targets for the APM would need to be increased in order to ensure that quality for APM patients increases to match what would be expected outside of the APM and that payments are adequate to cover the costs of the new technology.

Once the desired reduction in spending or improvement in quality has been achieved, the Target(s) for the APM would need to change to *maintaining* that lower spending level or improved level of quality. Maintaining spending or quality over time does not mean that the Target should remain exactly the same every year, for all the reasons described above – changes over time in the costs of services, the technologies available for diagnosis and treatment, evidence about appropriateness of services, etc. will make it more or less expensive to continue doing essentially the same thing and/or allow better outcomes than were previously possible.

If it is possible to achieve additional savings or additional improvements in quality, the performance Target(s) could be updated to reflect that, but a determination would also need to be made as to whether changes in other parameters were needed, such as payments for additional types of services or higher payment amounts for services.

2. Alternative Approaches to Updating Parameters

Two fundamentally different approaches can be used to update the parameters of an APM:

- An **analytic approach** that uses analyses of data about costs, outcomes, etc. in an effort to determine what the “right” changes in the APM parameters should be for all providers. Most of the material in Sections VI and VII describes analytical approaches to setting payment amounts and performance targets.
- A **competitive approach** that allows individual providers to determine the changes in the APM parameters for services they deliver based on the costs and outcomes they believe they can achieve, and then payers or patients choose providers based on the parameters they set.

Each of these methods, as well as combinations of the two methods, are used for updating payment amounts in current payment systems. For example:

- CMS uses a primarily analytic approach in making annual updates to physician payments under the Medicare Physician Fee Schedule. New CPT/HCPCS procedure codes are assigned a number of Relative Value Units (RVUs) based on an estimate of the relative amount of time and costs associated with performing the procedure/service compared to other procedures, and the RVUs for existing procedure codes are updated when information indicates that they are “misvalued” (e.g., because of changes in technology or improvements in efficiency). The conversion factor that translates RVUs into dollars is also updated based on estimated/projected increases in unit costs due to inflation and other factors. The same RVU and conversion factor are then used to pay all physicians in a particular geographic area when they perform a particular procedure or service.²⁸⁵
- For healthcare services that are typically not covered by insurance, such as Lasik surgery for vision correction and cosmetic surgery, the healthcare providers that deliver those services use a competitive approach to update the fees they charge patients. Different providers can charge different fees and can also commit to different outcomes. A provider that charges too much for a service without correspondingly better outcomes may find they have no patients if there are other providers willing to charge less for the same service with similar outcomes.
- A combination of the two approaches is commonly used to determine the amount that private health plans will pay physicians and hospitals for services that are covered by insurance. The Medicare RVU structure for physician fees or DRG structure for hospital payments is used to determine the relative amounts that will be paid for different services or patients, but the conversion factor that is applied to those RVUs or DRG weights is negotiated between each payer and provider. The size of the conversion factor depends on the level of competition among providers as well as the size of the payer.
- Medicare uses a combination of competitive and analytic approaches in determining payments for some

Durable Medical Equipment (DME) items. In larger metropolitan areas, Medicare selects providers of specific Durable Medical Equipment (DME) items through a competitive bidding process. Medicare then sets fees in smaller regions based on the payment amounts determined through the competitive bidding process used in the larger regions.²⁸⁶

a. Advantages and Disadvantages of the Alternative Approaches

There are advantages and disadvantages to both the analytic and competitive approaches. The significance of the advantages and disadvantages differs depending on (1) the types of patients, conditions, providers, and markets that are involved, and (2) whether patient-level or population-level performance targets are being used, so different approaches may be appropriate for different APMs and even for the same APM in different communities.

- Analytic approaches work best when most patients need similar numbers and types of services and the costs and outcomes that can be achieved for patients are similar, since that allows robust estimates of appropriate payment amounts, performance targets, etc. that can be used by all providers, including small providers, for the patients they are likely to treat. It is more difficult to accurately and reliably estimate appropriate parameters when patients vary significantly in terms of their needs and there is considerable variability in the outcomes that can be achieved. In contrast, a competitive approach can allow greater flexibility for providers to customize their approaches to address unique patient needs.
- Analytic approaches tend to reinforce the status quo, whereas competitive approaches allow and encourage innovative approaches to care delivery. For example, cost accounting systems can enable one to accurately determine the lowest amount that it currently costs to deliver a service today, but they cannot predict whether an innovative provider could redesign the service in a way that reduces costs even more. Similarly, quality benchmarking systems can determine the best level of quality that is being achieved today, but they cannot predict how much an innovative provider could improve quality by delivering care in a different way.
- Although competitive approaches can stimulate innovation, the innovations may or may not prove to be desirable. For example, a provider might offer to deliver care at a lower cost by using an approach that will turn out to cause more severe complications for patients in the long run, but that will not be apparent until after the services have already been delivered. In contrast, analytic approaches are more likely to support conservative approaches that are known to be feasible and sustainable with limited potential for unintended consequences.
- A competitive approach requires a mechanism for selecting one of the competing providers to deliver services to a particular patient. Competition at a population level (i.e., selecting providers based on their average cost and quality for a group of patients) can

make it more difficult for patients with unique needs to obtain the highest-quality, most affordable care. Competition at the patient level (i.e., allowing individual patients to select providers based on their cost and quality) is more likely to result in better results for individual patients, but individual patients may not always be able to make good choices among providers (e.g., in an emergency situation, when there are complex tradeoffs involved, or when the patient has cognitive limitations or difficulties communicating with providers).

- In many communities, there is only one provider that delivers a particular service (or only one that is willing and able to manage the care for patients with unique needs), either because the community is too small to support more than one provider or because all of the providers have consolidated into a single entity. If there is only one provider in a community, it is impossible to rely on a competitive approach in that community.
- In a competitive approach, a large provider organization will be better able to accept payments that are below cost in the short run than will small provider organizations. This “loss leader” approach could force the small provider organizations out of business, thereby reducing the ability to use a competitive approach in subsequent years.
- There may be legitimate reasons why one provider will incur higher costs to deliver the same service than another provider. For example, the average cost of delivering a service will likely be higher for a provider located in a small rural community than for a provider in an urban area simply because the volume of services will be lower and the unit costs of personnel and supplies may be higher. If patients who have the wherewithal to travel do so in order to get the lower price from higher-volume providers, the cost of the services for those who cannot travel will be even higher.

In healthcare, analytic approaches and population-level competition are used far more often than patient-level competition, whereas in other industries, the reverse is true. In most industries other than healthcare, the supplier of a product or service has the ability to determine how much they will charge an individual consumer and what performance standards the product or service will meet, and the consumer of the product or service decides whether to purchase the product or service on that basis. If a particular supplier is able to offer a particular product at the same quality but a lower price, they are free to do so and consumers can choose that supplier over others. If a particular supplier can offer a product or service at a higher level of quality but charges a higher price to do so, a consumer can decide whether to pay for the higher quality or to use a supplier that offers a standard product/service at the standard price.

b. Problems with Current Approaches to Encouraging Competition

There has been growing interest in finding ways to introduce more patient-level competition into healthcare in order to control healthcare costs while improving quality. The two most common approaches that have been pursued to date – “transparency” about prices and quality, and higher cost-sharing by patients – have had disappointing results, but this is likely due in part to the failure to make any changes to the underlying payment system:

- **Knowing the Prices of Individual Services Does Not Enable Patients to Determine the Amount They Will Pay for Their Care.** It is obviously impossible for patients to choose lower-priced providers if they don’t know what the provider charges for services. However, because most “price transparency” initiatives are implemented within a fee-for-service payment context, the patient will at best know what a provider charges for individual services, not the total amount the patient will need to pay for the full set of services the patient will actually receive. A patient could choose a provider that charges lower prices for individual services but the patient (or their insurance plan) could end up spending more if that provider uses more services or more expensive services than other providers that charge more for individual services but use fewer of them.
- **Historical Data on Total Cost of Care and Episode Spending Do Not Enable Patients to Determine How Much They Will Need to Spend on Their Care in the Future.** Because of the limitations of reporting only the prices of individual services, some transparency initiatives have begun reporting the total amount that has been spent on all of the services that patients received during a particular episode of care or during a particular period of time. However, these estimates are averages based on *past* performance for *other* patients, not a guarantee of what care for an individual patient will cost in the future.
- **Historical Data on Population-Level Quality Performance Does Not Assure Individual Patients of High-Quality Care.** A patient may not want to choose a lower-cost provider unless it is clear that the patient will receive equal or better-quality care than the patient would have received from a higher-cost provider. However, publicly-reported quality measures are calculated at a population level and rarely show that 100% of patients received high-quality care, so even if the lower-cost provider has better quality scores, an individual patient has no way of knowing whether the care they will receive will meet quality standards or not.
- **Higher Cost-Sharing May Discourage Use of Higher-Value Care.** One of the fundamental differences between health care and other industries is that the recipient of the service does not pay the full price of the service. Although high-deductible health plans and higher co-insurance rates have been promoted as ways of discouraging patients from using unnecessary services and making patients more sensitive to differences in the prices of necessary services, these

approaches also discourage patients from using desirable services and they do relatively little to encourage price-shopping for high-priced services.

c. Using APMs to Support Patient-Level Competition on Cost and Quality

These weaknesses can be overcome by combining a well-designed Alternative Payment Model with appropriate mechanisms for transparency and patient cost-sharing. For example, the condition-based bundled payments described in Section VI create greater certainty than fee-for-service payment about the total amount that will be spent on all of the services a provider or team of providers will deliver to an individual patient in order to treat or manage a particular health condition or combination of conditions. Patient-level accountability targets create greater certainty for individual patients as to the quality of care they will receive and the outcomes they will achieve than the population-level approach to quality used in most APMs. Patient cost-sharing can be based more directly on the differences in prices between services through techniques such as reference pricing discussed in Section VI.A.

Once an APM is developed that enables patients to easily and accurately compare the cost and quality of providers, the following combination of analytic and competitive approaches could be used to encourage innovation while protecting patients who do not have choices of providers or who are not able to make effective choices.

- 1. Set Default Parameters Using an Analytic Approach.** Initially, a payer would establish payment amounts and quality standards/targets for the APM using analytic methods. These would be considered the “default” parameters, i.e., the amounts that would be paid to a provider and the quality standards the provider would be expected to meet for patients with specific characteristics unless the provider explicitly proposed a different set of parameters.
- 2. Allow Individual Providers to Set Different Prices and Performance Targets.** An individual provider (or team of providers) that is participating in an APM could voluntarily choose to:
 - ♦ commit to meet higher quality performance targets for a particular category of patients at the default price, and to only accept payment when those standards or outcomes were achieved;
 - ♦ commit to achieve the default quality targets at a lower price; or
 - ♦ commit to meet higher quality performance targets for one or more categories of patients, but charge a higher price for doing so.

Each provider’s prices and performance targets would be posted publicly in a format enabling patients to easily compare providers.

A provider that wants to deviate from the default prices or performance Targets will presumably want to use some kind of analytic approach to determine what prices or performance Targets to use. However, the methodology used by the provider may be different from the methodology used by the payer to establish the default parameters, particularly since the provider may have access to different infor-

mation than the payer. For example, the provider will likely have a better understanding of what aspects of its costs could be reduced than any payer would, and the provider will likely have a better ability to determine which subgroups of patients could benefit from a different approach to care delivery.

- 3. Allow Patients to Choose Providers Based on Prices and Quality.** If the higher quality standards proposed by a provider were of value to patients, the patients could choose to receive care from those providers/teams rather than other providers/teams that only committed to meet the default standards and Targets. If the provider charged more for higher quality, the patient could be required to pay the difference between the provider’s charges and the default payment amounts in addition to any cost-sharing the patient would need to pay based on the default amounts. (If patients who could afford to pay the difference aren’t willing to pay for it, that could indicate that the “higher quality” care was not valued highly enough by the patients to justify delivering it.) If the provider charged less for the same level of quality, the patient’s default cost-sharing could be reduced by the difference between the provider’s charge and the default payment amount.
- 4. Update Default Parameters Based on Provider-Determined Prices and Quality Targets.** If patients demonstrated a preference for providers/teams that had committed to higher quality Targets or charged lower prices, other providers/teams would have an incentive to also meet the higher Targets or to charge the lower prices in order to retain patients and attract new patients. Once a majority of provider/teams have voluntarily committed to meet higher Targets or charge lower prices for a subcategory of patients, the higher Targets could then become the new default Targets for all APM providers. Some providers could then choose to adopt even higher performance Targets or to charge even less for care to once again to distinguish themselves from other providers, continuing the virtuous cycle of improvements in value.

Although this type of competitive model is rare in healthcare, the approach has been used successfully. In Minnesota in the 1990s, the Buyers Healthcare Action Group (BHCAG) developed the “Patient Choice” payment system. “Care systems” (groups of providers, but not necessarily integrated delivery systems) submitted bids on the risk-adjusted amount of payment they felt would be needed to provide all of the healthcare services patients would need. The care systems were divided into cost tiers based on their relative bids; consumers selected a care system based on both cost tiers and quality information, and the consumers paid the difference in the bid price if they selected a care system in a higher cost tier. Providers continued to bill based on CPT codes (with payments also authorized for some previously unreimbursed codes), but the payment rates for individual services were adjusted up or down to keep total payments within an overall spending target. The spending target was based on the provider’s bid, but it was adjusted upward or downward based on the relative illness and other characteristics of the patients that the provider actually cared for (this was intended to ensure

Example of a Competitive Approach to Cost and Quality Supported by an APM

In a hypothetical community where there are several different hospitals and orthopedic surgery groups that offer total knee replacement (TKR) surgery, the providers and purchasers could work together to define an APM for TKR. Under the APM:

- For low-risk patients, orthopedic surgery practices would receive a bundled/warranted payment to cover the full cost of the procedure, rehabilitation services, and treatment for any complications of the procedure or rehabilitation. The surgery practice would receive no payment at all if the patient fails to achieve a predefined outcome after rehabilitation has been completed (e.g., ability to walk for a specific distance without significant pain). Based on an analysis of current costs, the providers and purchasers all agree that TKR services can be delivered to low-risk patients for \$25,000, including costs associated with current rates of complications and readmissions. Data show that 95% of patients currently achieve the desired outcome. Consequently, the purchasers agree that they will pay \$26,315 for TKR to low-risk patients based on the expected cost and the expected outcomes ($\$26,315 = \$25,000/95\%$)
- For higher-risk patients, the surgeons, hospitals, and other providers will be paid fees for their individual services, but under the APM, the fee to the surgeon will be reduced if the total payments for TKR services for an individual patient exceeds a risk-stratified target amount for that patient or if the patient fails to achieve outcome targets deemed feasible based on their risk status.

Each surgery practice publicly posts its charge for performing TKR surgery on low-risk patients and also posts its complication rate, readmission rate, and rate of successful outcomes for those patients. Most surgery practices charge the agreed-upon rate of \$26,315, but two practices set a different charge: Surgery Practice A sets a rate of \$25,000 because it believes it can perform the procedure at lower cost. Practice B sets an optional rate of \$27,500 for patients who want to achieve better outcomes more quickly. Patients who are interested in the better/faster outcomes could choose Practice B but the patient would pay the extra \$1,185 charge themselves (i.e., the difference between the \$27,500 charge and the standard payment of \$26,315). Patients who choose Practice A would receive a rebate or bonus of \$657, 50% of the difference between the lower charge and the standard payment.

A significant number of patients choose Practice A because of the lower cost and because Practice A has shown that it achieves the same or better outcomes as the other surgery practices do. As a result, other surgery practices look for ways to reduce the cost of TKR so they can match the lower charge from Practice A. Once the majority of surgery practices are charging \$25,000, that becomes the new standard payment for TKR (with annual adjustments for inflation).

Each surgery practice also publicly posts its risk-stratified average spending and outcomes for higher-risk patients to help higher-risk patients decide which practice to choose. Surgery Practice C decides that instead of the standard APM methodology, it is willing to offer an outcome-based bundled/warranted payment to a subset of the higher-risk patients. It defines the characteristics of these patients and it publicly announces its willingness to provide TKR to these patients with the more accountable payment. A significant number of higher-risk patients begin using Practice C because of its commitment to a predictable payment and good outcomes. As a result, other surgery practices begin offering the same arrangement. Once the majority of surgeons do so, that payment approach becomes the new standard in the community for the subset of patients with those characteristics.

that providers were liable only for performance risk, not insurance risk associated with having sicker patients). Evaluations showed that the system encouraged patients to select more cost-effective providers and encouraged providers to reduce their costs while maintaining or improving quality.²⁸⁷

3. Issues in Updating APM Parameters

When analytic approaches are used, it may be possible to update APM parameters using the same methodology that was used for setting the initial values, but in many cases, a different methodology will be needed.

a. Updating Performance Targets Based on Comparison Groups and Counterfactuals

Several of the methodologies for defining population-level performance Targets described in Sections VI.B and VI.C are based in part on determining the spending or quality for a similar group of patients that are not

participating in the APM. As discussed in those sections, there are challenges in defining an appropriate comparison group for setting the initial parameters of the APM, particularly if the APM is focused on patients with particular characteristics on which data are not routinely collected.

The challenges in defining an appropriate comparison group will increase if the APM begins to be used more widely, because there will be fewer patients with the same characteristics who are not participating in the APM. Ultimately, if an evaluation determines that the APM is successful and it becomes the standard method of paying for a particular service or health condition, rather than merely an alternative to fee-for-service payment, it will be impossible to continue setting performance Targets using a comparison group.

Similar concerns arise with spending Targets defined based on counterfactuals, i.e., estimates of what the provider would have spent or what outcomes would have been achieved if the APM had not been imple-

mented. It will become increasingly difficult over time to accurately estimate “what spending (or quality) would have been” in the absence of the APM.

b. Updating Targets Based on Prior Performance

One of the alternatives to using comparison groups defined in Section VI is to set Targets based on the provider’s own performance prior to the initiation of the APM. If the Target Change from this baseline is defined to be zero (i.e., there should be no worsening of quality or no increase in spending from the prior performance level), then the initial Target could simply be updated regularly based on changes in unit costs or changes in evidence about appropriate quality standards.

However, if a non-zero Target Change is used initially (i.e., there is an expectation of some reduction in spending or improvement in quality from the baseline level), then once that Target is achieved, an assessment would need to be made as to whether further reductions in spending or quality are feasible. At some point, it will become impossible to continue reducing spending without harming quality (at least until new technologies are developed), and in these cases the Target Change would need to be redefined to be zero, since the goal would become maintaining the current level of performance rather than continuing to reduce it. The performance Target would then be updated regularly to reflect changes in unit costs, new evidence regarding appropriate treatment, etc.

c. Updating Spending vs. Utilization Targets

One of the primary reasons that performance Targets for spending need to be updated over time is because of changes in input prices, e.g., increases in employee wages, changes in the prices of drugs, devices, and other supplies, etc. Even if a provider is delivering or ordering the exact same services and achieving the exact same results year after year, spending will change over time due to changes in the unit costs of those services.

As discussed in Section VI, because changes in input prices are generally beyond the control of providers, it is often preferable to define performance Targets in terms of utilization measures rather than spending. If the performance Targets are defined in terms of spending, it will generally be necessary to update the Targets by splitting spending into utilization of services and the prices for services, updating each separately, and then recombining them into an updated spending target.

d. Eliminating Disparities in Payment Amounts and Targets

The creation of an Alternative Payment Model can reveal disparities in the amounts that are being spent for care and the outcomes that are being achieved for that spending that were not visible under the current payment system. For example, even if a payer is paying the exact same amount to every provider for each individual service they deliver under the current payment system, differences in the number and types of services providers use to treat their patients mean that the total spending to treat a patient will differ across providers and pa-

tients. If an APM is established that pays providers a single bundled payment to treat a patient, a common approach to determining the amount of that payment for each provider is to make it equal to the combined amounts for individual services that each provider is receiving under the current payment system. This approach ensures that individual providers do not lose revenue compared to the current payment system and that the payer does not spend any more than they are spending under the current payment system. However, the payer would then be paying different amounts to different providers for delivering the “same” services to similar patients.

If the APM does not bundle payments for services but creates accountability for the total amount spent on a particular treatment or condition, a target based on maintaining or reducing the provider’s current level of spending will result in different Targets for different providers. Similarly, if accountability for quality is based on maintaining or improving the provider’s current level of quality or outcomes, then the Targets will differ if providers differ in their current levels of quality or outcomes.

In this situation, the differences in payments and outcomes are not *created* by the APM; they already existed under the standard payment system. The APM simply made the differences *visible* by paying for the entire bundle of services at once rather than paying for one service at a time, or by creating specific accountability Targets for spending and quality.

It is well known that there is significant, unjustified practice variation across the country, within individual regions of the country, within individual communities, and even within the same physician practices or health systems. In some cases, the variation reflects the delivery of unnecessary services, and thereby represents an opportunity for reducing spending without harming outcomes or to improve outcomes without increasing spending. In other cases, the variation may reflect underuse of services or poor quality of care, and in these cases, higher spending may be needed. In these cases, it would be undesirable to allow the differences in payment amounts and targets to persist.

Before trying to equalize payment amounts or targets, however, it is essential to first determine whether there are legitimate reasons for the differences. For example, if one provider’s patients are different from another’s in some way that legitimately requires more services to be delivered to achieve good outcomes, the patients would be harmed if that provider were suddenly only paid the same amount as other providers are. The appropriate response is not to simply continue paying different amounts, but to revise the risk stratification/adjustment structure of the APM so that comparing payment amounts across providers is truly an “apples to apples” comparison. Similarly, if outcomes are different because some providers have patients who have a higher underlying risk of poor outcomes, then the risk stratification structure of the APM should be revised to address that.

When an APM is first implemented, it can be difficult to determine the extent to which differences in payment

amounts or targets represent disparities in performance or inadequate risk adjustment in the APM itself. Moreover, patients, providers, and/or payers may be unwilling to participate in the APM if the payment amounts or targets differ too dramatically from current levels of revenue/spending and performance. To address this, an APM can begin with customized payments amounts and targets for each provider that are based on the past performance levels of that provider, and then transition to payment amounts and targets that are common to all providers or all providers with similar characteristics. One approach to doing this is to use payment amounts and targets that are a weighted average of a provider-specific amount/target and a regional or national amount/target, and then change the weighting over time so it increasingly favors a common regional or national amount/target.

Example: Prior to creation of the Inpatient Prospective Payment System (IPPS) in 1983, hospitals had been paid by Medicare based on their actual costs. IPPS established uniform national payment amounts for each Diagnosis Related Group (DRG), but these amounts were phased in over a three year period using a weighted average of a hospital-specific payment rate, a regional payment rate, and the national payment rate. During the first nine months of implementation (January – September 1984), the DRG payments to an individual hospital were based 75% on a specific rate for that hospital based on its historical costs and 25% on the average rate for hospitals in the same geographic region. Two years later (in 1986), only 25% of the hospital's rate was based on its historical costs and 75% on the national rate. Beginning in 1987, 100% of a hospital's payment was based on the national rate.

Example: In the CMS Comprehensive Care for Joint Replacement Program, Target Prices are determined through a combination of the average amount that a provider has been paid in the past for a joint replacement episode and the average amount paid to all providers in the same geographic region. In the first two years of the program, the Target Price is a weighted average based two-thirds on the provider's past payments and one-third on the average payments made to providers in the region; in the third year, the weighting shifts to one-third based on the provider's past payments and two-thirds based on the regional average; and in the fourth and fifth years, the Target Prices are based entirely on the regional average.²⁸⁸

IX. CREATING BETTER VALUE-BASED PAYMENT MODELS

There are many opportunities for reducing healthcare spending without harming patients, and there are also many opportunities for improving the quality of care without increasing spending. Many of these opportunities exist because of problems created by fee-for-service payment systems, and alternative methods of payment are needed in order to successfully pursue these opportunities.

There is generally no one “best” way to design an APM for any individual improvement opportunity. As this report makes clear, there are multiple ways to structure payments and accountability mechanisms, each of which has strengths and weaknesses. However, this report also shows that there are better and worse ways of designing an APM for addressing any specific opportunity for improvement. Many current APMs have failed to achieve significant savings or improvements in quality because they do not correct the problems created by current fee-for-service payment systems. This is due to flaws in the design of *those specific APMs*, not to flaws in the basic concept of an Alternative Payment Model.

In addition, there are areas such as primary care where payments are not only structured poorly but where the amounts of payment are inadequate to cover the costs of high-quality care, and an APM may not be the best way to correct those problems.

A. Correcting the Problems with Fee for Service Payment

The tools described in this report show how APMs can be designed in ways that do correct the problems with fee-for-service payments. These APMs are more likely to be successful than current APMs. Specifically:

- **A well-designed APM pays for the high-value services needed to improve patient care.** To be successful, an APM must make any changes needed in the way providers are paid so they are able to deliver the services that will improve outcomes and reduce spending. Most current APMs do not make any changes in the ways that providers are paid, but merely provide “incentives” to reduce spending or improve quality.
- **A well-designed APM aligns the amount of payment with the cost of delivering good care.** An APM must change the amounts paid for individual services so payments are aligned with the actual costs of delivering services, regardless of the volume of services delivered. Many current APMs actually widen the gap between payments and costs rather than narrowing it.
- **A well-designed APM assures patients that they will receive appropriate, high-quality care that will achieve a good outcome for them (not just for other patients).** APMs can and should be designed with patient-level

quality standards and targets that tell each individual patient in advance what they can expect in terms of quality and outcomes. Most current APMs only assess whether quality has changed on average for a group of patients, not whether it has improved or worsened for individual patients.

- **A well-designed APM makes the cost of healthcare services more predictable and comparable.** APMs can and should specify in advance the amount that a provider will be paid and the total amount that will be spent for treatment of a particular condition or combination of conditions so that patients can compare the costs of care across providers. Many current APMs do not set spending targets until after services are already delivered, and most do not even make final determinations as to which patients are eligible for the APM until after services are delivered, making it impossible for a patient or their payer to know in advance how much will be spent on the patient’s care.

Well-designed APMs represent a radical shift from the way services are paid for under both current fee-for-service systems and current APMs. However, the differences only seem “radical” to those in the healthcare industry. Well-designed APM actually move healthcare payment closer to the way products and services are paid for in other industries. For example:

- In other industries, customers know the full price of the product or service before they buy it and they can compare the prices different suppliers charge for similar products and services. Under fee-for-service payments in healthcare, patients and payers cannot even obtain an estimate of the combined fees for all of the services they will receive in order to be treated for a health problem, much less a guaranteed price for an entire package of services. An APM that uses a bundled payment and/or specific spending targets creates more of the kind of certainty about spending that consumers have in other industries.
- In other industries, customers expect products and services to have a warranty against defects and a money-back guarantee of performance. Under fee-for-service payments in healthcare, physicians, hospitals, and other healthcare providers are paid for delivering services regardless of whether the services are delivered in the highest-quality way, regardless of whether the services have positive or negative effects on the patient, and regardless of whether the services were necessary or appropriate for the patient in the first place. An APM that incorporates warranties and performance guarantees rewards providers that deliver high-quality products and services, and well-designed APMs also encourage providers to clearly define the outcomes their services can and cannot be expected to achieve.

B. Preserving the Strengths of Fee-for-Service Payment

Many current APMs have had poor results not only because they fail to correct the problems with fee-for-service systems, but also because they fail to preserve its strengths.²⁸⁹ Some types of proposed APMs would be even worse in this respect. Although there are serious problems with the fee-for-service payment system, it would not have persisted for so long without any redeeming features. Well-designed APMs must also preserve four strengths of the fee-for-service payment system:

- **A provider is only paid if a patient receives a service.** Although there are clearly serious problems with the quality and cost of the services delivered under fee-for-service payment, the system at least gives patients and payers the confidence that they only pay something if they receive something in return. Under many “population-based payment” APMs, providers would be paid even if they do nothing for patients. In contrast, a well-designed APM will ensure that patients who need help receive it.
- **Payments are higher for patients who need more services.** Although fee-for-service payment is criticized for rewarding “volume over value,” any payment system that doesn’t adequately support a higher volume of services *when more services are needed* can result in worse outcomes for patients and higher long-run costs. Many APMs fail to adjust payments for im-

portant differences in patients that require more services or more expensive services.²⁹⁰

- **A provider’s payment is based on things the provider can control.** Although fee-for-service payment fails to hold providers accountable for problems they caused or could have prevented, it also does not penalize them for things outside of their control. Many current APMs go too far in the opposite direction – placing healthcare providers at financial risk for the total cost of care even though they can only control or influence a small part of it. In other industries, warranties and performance guarantees are typically limited to correcting defects the producer caused or could have prevented, and a good APM will do the same in healthcare.
- **A provider knows how much they will be paid before delivering a service.** Under fee-for-service payment, a provider knows exactly what they will be paid for delivering a service before they deliver that service, so the provider can determine whether they are likely to receive sufficient revenue to cover their costs before they incur those costs. Under many APMs, it is impossible for the participating providers to predict how much they will be paid for the services they will deliver, and they may not know for sure how much they will receive until many months after the services are actually delivered. In contrast, in a well-designed APM, payment amounts and Targets are clearly defined in advance, so providers, patients, and payers all know what will be paid and spent.

TABLE 17
COMPARISON OF WELL-DESIGNED APMs TO CURRENT APMs and FFS

	Current APMs		Well-Designed APMs			
	Shared Savings	Population-Based Payment	Accountable Payment for Services	Accountable Bundled Payment	Outcome-Based Payment	Bundled/Warrantied Payment
ADDRESSES WEAKNESSES IN FEE-FOR-SERVICE PAYMENT?						
Flexibility to deliver all needed high-value services?	NO	YES	YES	YES	YES	YES
Aligns payment with cost?	NO	NO	YES	YES	YES	YES
Assures each patient receives high-quality care?	NO	NO	YES	YES	YES	YES
Makes payments predictable and comparable?	NO	YES	YES	YES	YES	YES
PRESERVES STRENGTHS OF FEE-FOR-SERVICE PAYMENT?						
No payment unless a patient receives care?	YES	NO	YES	YES	YES	YES
Higher payments for patients who need more services?	YES	NO	YES	YES	YES	YES
Payment based only on things provider can control?	NO	NO	YES	YES	YES	YES
Provider knows payment before delivering services?	NO	YES	YES	YES	YES	YES

C. The Need for Multiple APMs

Patients have many different types of needs, and there are many different opportunities for reducing spending and improving quality. In Medicare, “fee for service” is not one payment system, but over a dozen different systems, each with multiple components and adjustments to address special circumstances.²⁹¹ This complexity exists because of differences in patient needs and differences in resources available in different communities.

For the same reasons, there is no one Alternative Payment Model that will be able to effectively support high-quality care to every type of patient or to effectively address all of the different opportunities for improvement. Multiple, different APMs will be needed. Although every effort should be made to design and operationalize these APMs similarly when possible, similarity of APM design primarily benefits payers, not patients or providers. In many cases, an APM that is trying to address different conditions will need to be different from other APMs in order to have the maximum impact on spending and provide the best assurance to patients that they will receive equal or better-quality care. Moreover, since so few different types of APMs have been implemented to date, there is no way to know which design is best, and a greater diversity of APM designs in the short run will help improve understanding of what works and what doesn't.

Many people erroneously believe that creating multiple APMs is undesirable because it will increase fragmentation of care and it will undercut efforts to improve coordination such as Accountable Care Organizations. However, if an APM is designed to encourage lower spending and better outcomes, then the providers participating in the APM will automatically have an incentive to address fragmentation problems and to improve coordination wherever that would truly achieve better results. The APM would encourage coordination when it is desirable, rather than trying to mandate arbitrary concepts of “coordination” that may increase costs without any corresponding benefits. As for ACOs, one of the biggest problems they have faced is that the shared savings APMs typically used to pay the ACO do not change the fee-for-service systems used to pay the individual providers who are part of the ACOs, making it impossible for those providers to pursue opportunities for savings and quality improvement. Well-designed APMs can help ACOs be more successful than they are today.

Whether one believes that patients will receive better care in an integrated delivery system, an ACO, from professional collaboration among independent providers, or through the patients' own choices of providers and self-coordination of services, the healthcare providers who are delivering care to the patients need to be compensated for their services. Even if a health insurance plan pays an integrated delivery system or multi-specialty physician group to address all of a patient's needs using a single capitated payment, that system or group will have to develop a way of compensating each individual physician, hospital, and other provider for what they do. If that compensation system uses a fee-for-service structure, as most such compensation systems do today, then it will create the same problems with patient care as would occur if the health plan made fee-for-service payments directly to each individual physician, hospital, etc.

“Population-based payment systems” do not solve the problems with fee-for-service payment, they simply shift them from payers to providers. In these situations, well-designed APMs will still be needed, but as methods of compensation for the individual providers.

D. The Need for Reforms Beyond APMs

While it should be a priority to find ways to reduce healthcare spending without harming the quality of care for patients and to improve the quality of care without increasing spending, there are also situations in which the only way to improve or even maintain current levels of quality will be to spend more than is being spent today. For example, many primary care practices, small rural hospitals, and safety-net providers are on the brink of closure due to either current payment systems that fail to adequately cover their costs or poorly-designed “value-based” payment models that have increased administrative burdens or shifted financial risk from large payers onto small providers.

Alternative Payment Models cannot solve these problems because the statutes authorizing APMs require that APMs reduce, or at least maintain, current levels of spending. Paying more to maintain access to primary care and preserve rural hospitals will help save lives, not reduce spending. While it is reasonable to ask primary care practices and rural hospitals to take greater accountability for the quality of care they deliver in return for receiving adequate payments, it is *not* reasonable to expect them to sustain themselves through “shared savings” on services they cannot control or to take “financial risk” without any financial reserves to fall back on. There clearly are many areas of avoidable spending in healthcare where savings could be produced to offset the higher payments needed to maintain safety net services, but the safety net providers are not the providers who are responsible for most of the avoidable spending, and the payments needed by safety net providers to sustain their services cannot be made contingent on whether other providers generate sufficient savings.

APMs are an important tool for improving quality and reducing spending, but other types of payment reform are also needed to sustain essential services, and those other payment reforms are needed soon, before it is too late to even preserve what currently exists.

E. Accelerating Progress Toward More Affordable Care

There is an urgent need to address the high and growing cost of healthcare in America and to do so in a way that improves the quality of care for citizens. Payment reform is a necessary part of the solution, but unfortunately, most of the payment reforms that have been pursued to date have had limited impact and, in some cases, have made things worse.

Alternative Payment Models and other types of payment reforms hold the potential for accelerating progress toward more affordable and higher-quality care if, but only if, they are designed in the right way. Faster progress in developing and implementing truly effective healthcare payment systems needs to be a national priority.

APPENDIX

Examples of Alternative Payment Models

There are many different opportunities for reducing spending and improving quality in healthcare, and no one Alternative Payment Model will be able to address all of them. However, there are also many similarities in the improvement opportunities across multiple types of health conditions, and so if an APM design effectively addresses an improvement opportunity for one type of health condition or patient, there is a good chance that a similar design could be effective for other types of health conditions and patients. Because different services are required to address different types of health conditions, the APM for each condition would have to be customized to that condition, but the general designs of the APMs, and the administrative systems needed to implement them, could be similar. For example:

- For many types of acute conditions, such as chest pain, childbirth, back pain, and others, some patients are receiving expensive, invasive procedures when less expensive, lower-risk treatments would be equally effective or more effective for the patient at a lower cost for the patient and payer. The specific procedures and treatment options differ significantly depending on the condition, but the barriers to improvement created by the current payment system are similar, and so the design of an APM to overcome those barriers could also be similar.
- For many different chronic conditions, such as asthma, COPD, diabetes, heart failure, and inflammatory bowel disease, many patients are hospitalized for exacerbations of their disease that could have been avoided through delivery of different treatments and other services than they receive today. The specific treatments and services will differ for each condition, but the barriers to delivering those services in the current payment system are similar, and so the design of an APM to overcome those barriers could also be similar.

For many types of health conditions, there are multiple opportunities for reducing spending and improving quality. For example, spending is higher than necessary for some chronic conditions due to both overuse of expensive drugs and other treatments and high rates of avoidable hospitalizations. Although the maximum impacts on affordability and outcomes would likely be achieved by pursuing every opportunity for improvement for a particular health condition, an APM designed to address multiple payment barriers will likely be more complex than one that is focused on correcting the barriers for a single opportunity for improvement, and the more complex APM may be more difficult for payers or small providers to implement. If a simpler APM that is focused on one opportunity can be implemented more quickly and more broadly, it may have a larger impact, at least in the short run, than an APM that tries to do many things.

To illustrate these concepts, and to illustrate how all of the components required for an APM can be constructed using the options described in this report, three different examples of APMs are described below. These examples are not intended to represent the full range of APMs that are needed to address all opportunities for improvements in healthcare, nor are they intended to represent the “ideal” APMs for the opportunities they are designed to address. Different options for some or all of the components will be more feasible, desirable, or effective for different types of providers and payers in different communities. The examples are:

- **Payment for a High-Value Service: Care Management for Chronic Diseases.** This example illustrates how an APM could make one targeted change to payment to address a specific opportunity for improvement, while leaving the rest of the payment system unchanged.
- **Condition-Based Payment for an Acute Condition: Maternity Care.** This example illustrates how an APM could replace the current fee-for-service payment system for the key services needed to provide care for a temporary condition.
- **Condition-Based Payment for a Chronic Condition.** This example illustrates how an APM could replace the current fee-for-service payment system for the many different kinds of services needed to diagnose and manage many types of chronic conditions.

All of the APM examples will be discussed assuming that the current payment system is structured the same as the payment systems used in the traditional Medicare program. The basic concepts described will be applicable to the majority of commercial payers, since most payers use payment systems similar to Medicare payment systems for most providers and most services, although they may pay different amounts for services, have different rules about which services are covered and when, etc. Adjustments to the APM designs might be needed for a payer that uses a different approach for paying for services that are a key element of the APM (e.g., for payers that pay for hospital inpatient stays on a per diem or percent-of-charges basis).

APM #1: Payment for a High-Value Service CARE MANAGEMENT FOR CHRONIC DISEASE

OVERVIEW OF THE APM

Under this APM, an individual who has been diagnosed with a chronic disease would choose a Chronic Care Management Team that is participating in the APM to provide care management services for one or more of the patient's chronic conditions. The patient would be classified into one of four need/risk categories based on characteristics that affect their likelihood of exacerbations and hospitalizations and the intensity of care management services the patient would need to prevent exacerbations and hospitalizations.

The Chronic Care Management Team would receive a quarterly Care Management Payment in addition to any fee-for-service payments the Team received for office visits, procedures, etc. needed to treat the patient's conditions. The amount of the Care Management Payment would be higher for a patient in a higher need/risk category. Except for patients in the Very High Risk category, the Team would not receive a quarterly Care Management Payment if the patient was admitted to the hospital during the quarter for reasons related to the chronic conditions the Team is supposed to be managing. For Very High Risk patients, the Team would be expected to maintain or reduce the rate at which the patients were being hospitalized before receiving the care management services.

The APM would reduce spending and improve outcomes by reducing the rate of avoidable hospital admissions.

DETAILS OF THE APM

1. Opportunity for Savings and Quality Improvement

Many patients with a chronic illness are admitted to the hospital one or more times during the course of a year because the symptoms of their illness become uncontrolled and sufficiently severe that they require inpatient treatment. This occurs with many different types of chronic conditions, including asthma, chronic obstructive pulmonary disease (COPD), diabetes, heart failure, and inflammatory bowel disease. For example, a patient with emphysema (one form of COPD) who does not use long-acting bronchodilators properly could develop severe difficulty breathing and require treatment with oxygen and medications in a hospital.

Each of these unplanned hospital admissions is expensive for both the patient and their health insurance plan. In addition, the patient may develop additional health problems during their hospital stay (e.g., a hospital-acquired infection), and if the patient is employed, they will miss work for several days. Reducing the likelihood

and frequency of these hospital admissions could generate significant savings for payers and achieve better outcomes for the patients.

2. Changes in Care Delivery Needed and Associated Costs

a. New and Different Services to Be Delivered

A variety of demonstration projects have shown that a large percentage of hospital admissions for exacerbations of a chronic disease can be avoided if a physician practice that is treating patients for the disease provides additional services to the patients. These services include:

- additional education to the patient about the situations that can cause exacerbations in their chronic illness and about steps that the patient can take to prevent these situations, training for the patient in how to use medications or other treatments, and education about the actions the patient should take to minimize the severity of symptoms when problems occur;
- visits to the patient's home to identify any factors that could make exacerbations more likely and help the patient correct those factors;
- regular contacts with the patient by phone, email, or other means to identify signs that their condition may be worsening and to make any appropriate changes in medications or other treatments;
- rapid response when it is determined that a patient's condition is worsening so that it can be treated without hospitalization whenever possible.

These services are generally referred to as "care management" services, since they do not involve treatment of the disease *per se*, but rather a set of complementary activities designed to improve the outcomes of treatment.

In most cases, it will be more efficient and effective to have a nurse or a trained community health worker deliver most of these care management services rather than a physician or other clinician. The patient's primary care provider or a specialist will have to determine whether changes in medications or other treatments are needed when the patient's condition worsens, but nurses, educators, and community health workers can provide most or all of the other services.

b. Cost of Delivering the New Services

The exact costs of delivering the care management services will vary depending on the type of staff used, the salary/wages needed to recruit and retain such staff in the community where they will work, the methods used for contacting patients, the time clinicians need to spend overseeing the new services and responding to issues identified by the care management staff, and other factors. However, the major driver of the costs will likely be the number of personnel needed to provide effective care management services. This will depend on:

- the number of patients in the practice who have the condition and the caseload that an individual staff member can effectively manage. If there are more patients than one nurse (or other type of staff) can effectively manage, then an additional staff member will be needed;
- the amount of time each patient will need from the staff to effectively address the patient's needs. This will depend on factors such as the patient's health literacy, the type of insurance coverage they have for medications, the relative severity of the patient's condition, and the presence of other specific comorbidities that affect the patient's likelihood of having an exacerbation of the chronic condition. Not every other health problem the patient has will have a significant impact on their ability to manage the chronic condition, but some (e.g., depression) can have a very significant effect. If a practice has a high percentage of patients with characteristics that will require more care management time, the maximum caseload that a staff member can handle will be smaller and the cost per patient will be higher.

There will also be startup costs involved when the services first begin. The new staff will need to be recruited and trained before they can deliver any services. Initial caseloads may be lower while patients are first enrolling in the service.

c. The Business Case for an Alternative Payment Model

An APM will be feasible for a particular chronic condition if analyses show that the expected savings from reduced rates of hospital admissions would be larger than the expected costs of delivering the care management services to patients who have that chronic condition. The analysis for a specific chronic condition would be based on:

- Estimates of the patient caseload that a nurse or other type of staff member can manage with the typical mix of characteristics of the patients who have the condition. These caseload estimates could be derived from the experience of demonstration projects in which similar services were delivered.
- The number of patients in a physician practice who have the condition. This would be based on medical records data from practices that would potentially participate and/or claims data from payers.
- The estimated cost of employing the staff and providing space, equipment, etc. to support their work. This

would be based on current labor market data and the experience of demonstration projects in which similar services were delivered.

- The rate at which patients with the chronic condition are currently being admitted to the hospital for exacerbations of the condition. This information could be obtained from healthcare claims data.
- Estimates of the reduced rate at which patients with the chronic condition who are receiving the care management services would be admitted to the hospital for exacerbations of the chronic disease. These hospitalization rate estimates could also be derived from the experience of demonstration projects.

3. Barriers in the Current Payment System

In general, under the current fee-for-service system, a primary care or specialty physician practice can only bill for face-to-face visits between a physician or other clinician (a nurse practitioner or physician assistant) and the patient. Assistance delivered through a phone call or email are generally not separately billable, and there is generally no payment for services delivered by a nurse, educator, or community health worker unless it is under the direct supervision of a physician or other clinician.

In recent years, several new billing codes have been added to the physician fee schedule by Medicare and other payers that allow physician practices to be paid for certain kinds of care management services in certain circumstances. However, the structure of these billing codes creates barriers to implementing care management services in the most efficient and effective way. For example, the Medicare Chronic Care Management Services payment is limited to patients with two or more chronic conditions, and the physician practice is required to document that it has provided at least 20 minutes of services to each patient each month in order to bill for the code.

Creating an APM would be appropriate if an analysis performed from the perspective of a physician practice showed that the revenue the practice could expect to receive from billing for services under the current payment system would be less than the costs of delivering the care management services in a way that would be expected to deliver the results assumed in the business case for payers.

EXAMPLE OF THE BUSINESS CASE FOR CARE MANAGEMENT FOR A CHRONIC DISEASE

The table shows a hypothetical example of 1,000 patients with a chronic condition. On average, about 13% of these patients are hospitalized during the course of the year for exacerbations of their condition. Stratification of the patients into four different need/risk categories shows that the rates of hospitalization vary significantly among the subgroups. Providers plan to hire registered nurses to provide education and self-management support to the patients; nurses who work with higher need/risk patients will have smaller caseloads. The provider expects to be able to reduce the overall rate of hospitalization by 50%, with different levels of reduction within each patient subgroup. The projected cost of the care management service is less than the expected savings from avoided hospitalizations, so paying to support the care management service would reduce total spending by 4%.

	<u>Current FFS</u>	<u>APM</u>	
Patients			
Total # of Patients	1,000	1,000	
Low Need/Risk Patients	450	450	
% Low Need/Risk Patients Hospitalized	5%	2%	
Medium Need/Risk Patients	325	325	
% Medium Need/Risk Patients Hospitalized	15%	6%	
High Need/Risk Patients	200	200	
% High Need/Risk Patients Hospitalized	25%	15%	
Very High Need/Risk Patients	25	25	
% Very High Need/Risk Patients Hospitalized	50%	30%	
Hospitalizations			
Total Hospitalizations	134	66	
Spending on Hospitalizations (\$10,000 per admission)	\$1,337,500	<u>\$660,000</u>	
Change in Spending on Hospitalizations		\$677,500 -51%	
Staffing for Care Management Service			
1 RN per 500 Low Need/Risk Patients		0.9	\$21
1 RN per 250 Medium Need/Risk Patients		1.3	\$42
1 RN per 100 High Need/Risk Patients		2	\$104
1 RN per 25 Very High Need Patients		<u>1</u>	<u>\$417</u>
Total RNs		5	
Cost of Care Management Services (\$100K/RN + 25% overhead)		<u>\$625,000</u>	
Total Spending on Hospitalizations and Care Management	<u>\$1,337,500</u>	<u>\$1,285,000</u>	
Savings		\$52,500 -4%	

4. Design of the APM

The lack of adequate payment in the current fee-for-service system could be addressed by modifying the current billing codes or creating one or more new billing codes specifically designed to support the desired care management services. However, simply modifying the fee schedule would provide no assurance to the patient or payer that the services paid for would achieve the desired results.

The Alternative Payment Model described below can enable payments to be made for care management in a way that ensures patients will benefit and spending will decrease:

a. Defining the Eligible Patients and Physician Practices

i. Eligibility Criteria for Patients

Patients would be eligible to receive services supported by the APM if:

- (1) they have a “chronic condition” consisting of one or more chronic diseases that are included on a specified list of chronic diseases, and
- (2) either
 - (a) their condition has reached a specified level of severity and/or
 - (b) they have other characteristics increasing the risk of hospitalization.

Eligibility could be based on different combinations of minimum severity levels and other characteristics in order to focus on all patients whose risk of exacerbation-related hospitalizations would be sufficiently high to justify delivering care management services.

ii. Designation of the Chronic Care Management Team by the Patient

A Chronic Care Management Team (consisting of the physicians or clinicians in a single practice, clinicians from two or more collaborating practices with different specialties, or clinicians employed by a hospital or health system) would be eligible to participate in the APM if the Chronic Care Management Team treats patients with the chronic condition and is willing to deliver care management services supported by the APM to eligible patients.

If an eligible patient wanted to receive the enhanced services and avoid hospitalizations, the patient would designate one of the Chronic Care Management Teams participating in the APM to manage their care for the condition for a period of at least three months. The patient could change this designation at the beginning of any calendar quarter. Only one Chronic Care Management Team could receive payments under the APM for an individual patient during each three-month period. For patients who have multiple chronic diseases that require coordinated management, the Chronic Care Management Team would be responsible for providing high-quality, coordinated care management services for

all of the diseases (i.e. for the patient as a whole), rather than any individual disease.

Before the patient designated the Chronic Care Management Team to manage their care, the Team would describe the care management services it would commit to deliver to the patient. Before accepting the patient, the Team could ask the patient to commit to actions that would support good outcomes (e.g., taking prescribed medications, contacting the practice when a problem arises, etc.).

In order to receive payments from a patient’s health insurance plan, the Chronic Care Management Team would need to verify and document that the patient met the eligibility criteria and document that the patient had designated the Team to manage their care.

b. Removing the Barriers in the Current Payment System

i. Stratification of Patients Based on Need/Risk

For each chronic condition, criteria would be defined for stratifying patients into four categories of need and risk: (1) Low Need/Risk, (2) Moderate Need/Risk, (3) High Need/Risk, and (4) Very High Risk. The criteria would be based on characteristics of patients that are expected to affect the patient’s likelihood of exacerbations and hospitalizations and the intensity of care management services the patient would need in order to prevent exacerbations and hospitalizations. The specific criteria would differ for each condition; for example, the measure of the severity of disease would inherently be specific to each disease. The “Very High Risk” category would be limited to patients who have unique characteristics creating a high degree of uncertainty as to whether and how hospital admissions can be avoided.

If an eligible patient has selected a participating Chronic Care Management Team to manage their care, the Team would evaluate the patient and classify the patient into one of the four need/risk categories. A patient’s category could change from one calendar quarter to the next if the patient’s characteristics change (e.g., if the severity of their disease increases).

ii. Quarterly Payments for Care Management Services

The Chronic Care Management Team would be eligible to receive a standard, pre-defined Care Management Payment for each eligible patient on a quarterly basis from the patient’s health insurance plan, in addition to any other fee-for-service payments the Team received for office visits, procedures, etc. The Team could use the revenues from the Care Management Payments to pay for the nurses or other staff, equipment, travel, etc. needed to deliver the care management services. The Team would have flexibility regarding the exact services it delivered, the type of personnel it used, etc.

The amount of the Care Management Payment for a patient in a particular category would be based on the estimated average per-patient cost a Team would incur to deliver care management services to patients in that category, with adjustments for performance as de-

scribed below. A higher amount would be paid to the Chronic Care Management Team for patients in the higher need/risk categories. (These payments would be the type of stratified, bundled, condition-based payments described in Option 6 in Section VI.A, i.e., the payments would differ based on the patient's characteristics, rather than on the services provided, and a provider receiving the payment would have the flexibility to deliver multiple services and combinations of services.)

The amounts of the Care Management Payments would be periodically adjusted based on analyses of the actual costs incurred by Chronic Care Management Teams that successfully achieve the performance targets.

iii. Patient Cost-Sharing

The patient would not be required to pay any portion of the standard Care Management Payment. The patient would continue to pay cost-sharing for other separately billable services they received.

If it wished to do so, a Chronic Care Management Team would be permitted to charge more than the standard Care Management Payment amount for one or more categories of patients, and if the patient chose to use that Team, the patient would need to pay the difference between the Team's charge and the standard payment amount.

c. Creating Accountability for Utilization and Spending

i. Measures of Utilization/Spending

Two measures of utilization and spending would be used:

- **Rate of Condition-Related Admissions.** A participating Chronic Care Management Team would be accountable for how often its patients are hospitalized for exacerbations of the chronic condition. A definition would be developed as to which hospital admissions would be considered as related to the condition and which would not. If the "chronic condition" is defined as two or more chronic diseases, the condition-related admissions would include any admissions related to any of those chronic diseases, but not to other chronic diseases or acute conditions. The Team would only be accountable for the *rate* of condition-related admissions, not the amount of *spending* on the admissions, because the Team would not be able to control changes in the amounts paid to the hospital for an admission nor would it be able to control what happened to the patient after the patient was hospitalized that could lead to higher hospital spending.
- **Total Spending on Condition-Related Services.** In addition, a participating payer would measure the total amount of spending on all services that participating patients received that were related to the condition being managed under the APM. The Chronic Care Management Team would not be accountable for total spending, but the payer would monitor the total spending measure to determine whether the APM was increasing or reducing total spending related to the condition.

ii. Target Performance Rates for Condition-Related Hospital Admissions

Calculation of Benchmarks

- **Patients in the Low, Moderate, and High-Risk categories.** For patients in each of these categories, the national rate of hospital admissions for exacerbations of the condition would be calculated or estimated for similar patients during the year prior to implementation of the APM. Estimates would be made if data are not available during that year for all of the criteria needed to assign patients to the four categories. These rates would be the Benchmarks for the Chronic Care Management Teams participating in the APM. (This would be a "prior performance for similar patients" benchmark as described in the first alternative in Section VI.B.2.c.)
- **Patients in the Very High Risk category.** For the specific patients in the Very High Risk category who are being managed by the Chronic Care Management Team during the quarter, the proportion of those patients who were admitted to the hospital during the prior 1-2 years would be calculated and used as the Benchmark for that category for those patients. If prior years' data are not available for the patients, if their condition is newly diagnosed, or if the factors leading to their classification are new, then they will be assumed to have a high baseline rate of admissions based on their high-risk characteristics. (This would be a "prior performance for the same patients" benchmark.)

Determination of Target Changes

A Target Change would be defined for each of the four categories, based on the reduction in the rate of hospital admissions needed to offset the estimated spending on Care Management Payments for the patients in those categories. (This would be the "minimum change needed for success" as described in Section VI.B.2.d.)

Calculation of Target Rates

The combination of the Benchmark and the Target Change would define the Target for each category, i.e., the maximum percentage of patients expected to be hospitalized each quarter when care management services are being delivered. The Targets would be prospective, i.e., they would be defined before the participating Teams began delivering services. Except for the Very High Risk patients, the Targets for all participating Teams would be the same. Once the initial Targets were set, they could remain unchanged until such time as there was evidence that it was feasible to achieve lower rates or that the current rates could not be achieved by most Chronic Care Management Teams.

iii. Accountability for Hospital Admissions (Patients Not Classified as "Very High Risk")

Outcome-Based Payment Amount

The Chronic Care Management Team would only receive a Care Management Payment for a patient during a calendar quarter if that patient had no condition-

related hospital admission during that quarter. If the patient was hospitalized during the quarter for a reason related to the chronic condition being managed, the Team would receive no payment for the care management services delivered to that patient during that quarter.

The amount of the Care Management Payment the Chronic Care Management Team would receive for a patient classified into a particular category would be calculated by (a) estimating the per-patient cost of care management services for patients in that category, and (b) increasing that estimate by the Target percentage of patients not admitted to the hospital. For example, if the estimated cost of delivering services to patients in the moderate risk category was \$160 per patient per quarter, and the target rate of hospital admissions was 20%, then the payment per patient would be $\$160 / (1 - .20) = \200 . Under this approach, if the Chronic Care Management Team achieved the Target rate of admissions for patients in a category, then the Care Management Payments would match the expected costs of the care management services. If the actual rate of admissions is higher than the Target, the revenue from the Care Management Payments will fall short of the expected costs of the care management services, creating a financial penalty for the Chronic Care Management Team. If the rate of admissions is lower than the Target, the revenue from the payments will exceed the expected costs, creating a financial bonus for the Team.

Penalty for Failure to Impact the Performance Measure

After the initial year, if the rates of condition-related admissions in two or more categories are higher than the Target by a statistically significant amount when averaged over the previous 18 months, the Chronic Care Management Team would no longer be eligible to participate in the APM. The standard of statistical significance would be set at a level that balanced the expected rate of Type I and Type II errors based on the number of patients participating and the diversity of patient characteristics that affect the risk of hospitalization.

iv. Accountability for Hospital Admissions (Patients Classified as “Very High Risk”)

Base Payment Amount

The amount of the Care Management Payment a Chronic Care Management Team would receive for a patient in the Very High Risk category would be equal to the estimated per-patient cost the Team would incur to deliver care management services for patients in that category.

Penalty/Bonus for Performance

If the rate at which patients in the Very High Risk category were admitted to the hospital for exacerbations of the condition exceeded the Target by a statistically significant amount, the Chronic Care Management Team would repay 10% of the payments it had received for all patients in that category. If the admission rate was lower than the Target Rate by an amount that was both large in magnitude and statistically significant, the Team

would be paid an additional 10% for all patients in the category. The standard of statistical significance would be set at a level that balanced the expected rates of Type I and Type II errors. For Teams with a large number of Very High Risk patients, these penalties/bonuses could be calculated and paid quarterly; for Teams with smaller numbers of patients, the penalties/bonuses would be calculated annually.

v. Accountability for Total Spending

The payer’s average total spending per patient per month on all services related to the condition would be calculated for the year prior to initiation of the APM. The average total spending per patient per month for patients participating in the APM would be calculated for the second year that the APM is in operation. If the average spending for patients in the APM is higher than the average spending prior to initiation of the APM, the payer could choose to terminate or modify the APM.

d. Creating Accountability for Quality

i. Measures of Quality/Outcomes

Two quality measures would be used:

- **Rate of Condition-Based Admissions.** The primary measure of quality would be the utilization measure used for spending accountability, i.e., the rate of hospital admissions for exacerbations of the chronic condition. In general, patients will have a higher quality of life if they do not have exacerbations requiring a hospital admission and if they do not have to be hospitalized.
- **Mortality.** In order to ensure that hospitalizations are not being reduced by discouraging patients from being hospitalized when they need to be, a secondary measure of quality would be the rate of death among the participating patients during the months in which care management services are being delivered.

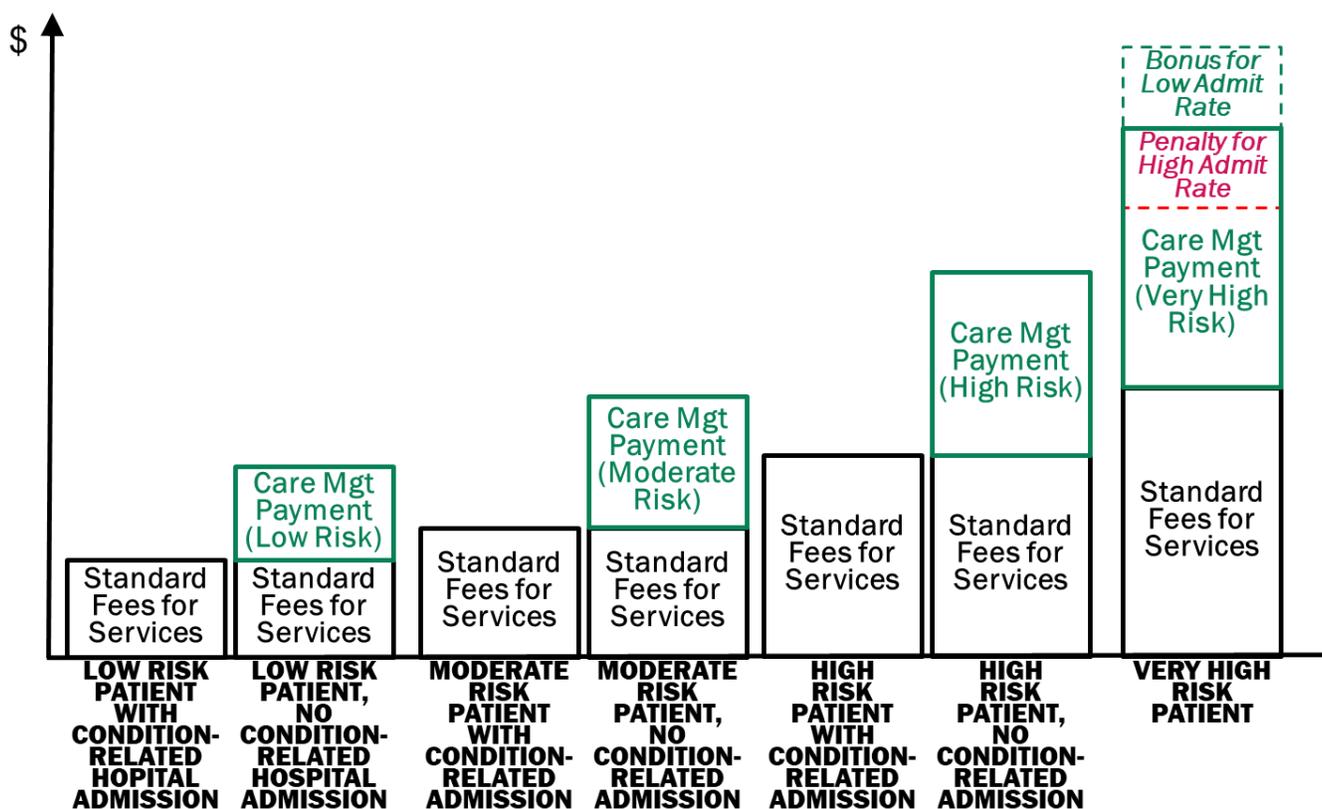
ii. Target for Mortality Rate

The rate of mortality for patients in each of the categories would be calculated for one or more years prior to the initiation of the APM. This rate would serve as the Target for that category under the APM.

iii. Accountability for Mortality

If the rate of mortality in one or more categories increased by a statistically significant amount for the patients of a Chronic Care Management Team that is participating in the APM, that Team would no longer be permitted to participate in the APM.

APM PAYMENTS AND ACCOUNTABILITY TO SUPPORT CARE MANAGEMENT FOR A CHRONIC DISEASE



5. Operationalizing the APM

In order for a Chronic Care Management Team to receive Care Management Payments under the APM, the team would submit a claim form on a quarterly basis for each eligible patient using one of four new codes:

- CM001 for a patient who meets the criteria for the Low Need/Risk category;
- CM002 for a patient who meets the criteria for the Moderate Need/Risk category;
- CM003 for a patient who meets the criteria for the High Need/Risk category; and
- CM004 for a patient who meets the criteria for the Very High Risk category.

The date of service on the claim would be the last day of the three-month period of care management services for which payment was being requested. The principal diagnosis code submitted with the billing code would be the condition being managed by the team under the APM. If the patient's eligibility was based on having a combination of two or more diseases, the principal diagnosis code would reflect the most severe disease and secondary codes would confirm the presence of the other diseases required for eligibility.

Submission of a claim form for a patient with one of these billing codes would represent a certification by the Chronic Care Management Team that:

- The patient met the eligibility criteria for the APM and for the assigned Need/Risk category.
- The Team had delivered three months of care management services to the patient.
- For codes CM001, CM002, and CM003, the patient had not been hospitalized during the three-month period for an exacerbation of the condition on which their eligibility was based. If the patient had been hospitalized for a different reason, the Chronic Care Management Team would document why the hospitalization was unrelated to the condition that the Team is managing.

If the patient's health insurance plan received a claim with a CM001, CM002, or CM003 billing code, before paying the claim, the plan would verify that the patient had not been hospitalized for an exacerbation of the condition being managed during the 90 days prior to the date of service recorded on the claim. If a hospitalization had occurred, the insurance plan could request additional documentation from the Chronic Care Management Team as to why the hospitalization was not related to the condition being managed.

A claim with a CM004 billing code would be paid regardless of whether the patient had been hospitalized during the previous quarter.

If the Chronic Care Management Team wished to charge patients more for care management services than the amounts that would be paid by their health insurance plans, the team would publish its charge for each of the four billing codes, and the patient would agree to the charge at the time that the patient was enrolling to receive care management services from the Team. The amount charged by a particular Team would be the same for all patients, regardless of their health insurance plan, and the Team would bill the patient for the difference between the charge and the amount paid by the plan.

At the end of each quarter, the Chronic Care Management Team would calculate the rates at which its patients had been admitted to the hospital for exacerbations of their condition. These rates would be calculated separately for each of the 4 categories and compared to the Targets for those categories along with a calculation of the statistical significance of the difference. Those comparisons would be provided to the Team's patients and to the health insurance plans for those patients.

The Chronic Care Management Team would make information on its condition-related hospital admission rates, the comparisons of the rates to the Target Rates, and its charges for care management publicly available so that patients could compare the cost and performance of different teams that manage that condition.

6. Implementing the APM

a. Obtaining Participation by Payers, Providers, and Patients

The APM would have a number of advantages for payers, providers, and patients that should encourage payers to implement the APM, encourage providers to participate in the APM, and encourage eligible patients to seek care from providers who are participating in the APM.

i. Advantages for Payers

- Participating health insurance plans could reduce spending on avoidable hospitalizations for plan members who have one or more types of chronic conditions.
- Health insurance plans could implement the APM by creating four new billing codes in their existing claims payment system.

ii. Advantages for Providers

- Participating Chronic Care Management Teams would receive additional payments that cover the cost of delivering the kinds of care management services needed to help their patients better manage their chronic conditions and avoid severe exacerbations that require hospitalizations.
- Participating Chronic Care Management Teams would receive higher payments to cover the higher costs of

providing care management services to patients with greater needs.

- Participating Chronic Care Management Teams would have the flexibility to deliver care management services in the ways that are most feasible for the providers on the Team and most effective for their patients.
- Participating Chronic Care Management Teams would only be held accountable for whether a patient they had explicitly enrolled for services was hospitalized for an exacerbation of the chronic condition the Team had committed to manage, not for the costs of the hospitalization or for other services the patient is receiving from the members of the Team or from other providers. The Team would know in advance what rate of hospitalizations it would be expected to achieve for its patients.
- Participating Chronic Care Management Teams would know when to expect payment and how much to expect based on the bills they submit to payers and the cost-sharing charged to patients. The largest financial loss the Team could experience would be the loss of the payments under the APM.
- Participating Chronic Care Management Teams could bill for services using their standard billing systems.

iii. Advantages for Patients

- Patients would have the choice of whether to receive the services, based on a clear understanding of the services they would receive, the actions they would need to take, and the results they could expect to achieve.
- Patients would know that the providers on their Team would be rewarded for helping them avoid exacerbations of their diseases but would have no financial incentive to withhold needed care.
- Patients would know how much they would need to pay for the services before choosing to receive them.
- Participating patients would experience fewer severe symptoms from their chronic disease. They would receive more care at home and require fewer visits to emergency departments and fewer admissions to hospitals to treat severe symptoms.

b. Finalizing the APM Parameters

A "beta test" of the APM will likely be needed with willing providers in order to finalize several key parameters of the APM:

- **Criteria defining the four categories of need/risk.** The categories should be defined so that they distinguish which patients will be at higher risk of exacerbations and which patients will need more care management services in order to avoid hospitalizations. However, data may not be available on all of the factors that would be expected to affect need and risk, and the APM will need to be implemented in order to enable those data to be collected.
- **Dollar amounts of the Care Management Payments.** The payment amounts should be based on the cost

the Chronic Care Management Team would incur in delivering the services, but the cost of the services will depend on the sizes of the patient caseloads that care management staff can support and the number of patients in each of the need/risk categories, and this can only be estimated after the services are actually implemented with support from the APM.

- **Benchmark rates of condition-related hospital admissions.** The performance Target and payment amounts will depend on the benchmark (baseline) rates of hospital admissions in each need/risk category, but this can only be determined after actual patients are classified into the need/risk categories.

Best estimates of these parameters would be used to initiate the beta test process, and the participating Teams would gather and share data from their actual experience in implementing care changes with payments under the APM in order to adjust the parameters.

APM #2: Treatment of an Acute Condition

MATERNITY CARE

OVERVIEW OF THE APM

A pregnant woman could choose a Maternity Care Team that is participating in the APM to deliver maternity-related services prior to, during, and following delivery of the baby. The Team would include all of the clinicians and providers needed to deliver the full range of care the woman could need, and the Team would ideally include at least one birth center as well as a hospital. The woman could change the Maternity Care Team at any time prior to the beginning of labor or during the post-partum period.

Under the APM, the Maternity Care Team would receive five different types of payments during the different phases of care:

- Monthly bundled payments for all pregnancy-related services needed prior to childbirth;
- A standby capacity payment for hospitals in the community to support the minimum capacity needed to offer labor and delivery services on a round-the-clock basis, particularly for high-risk pregnancies;
- A bundled/warranted payment for labor and delivery services, regardless of whether the delivery occurs in a birth center or a hospital;
- Monthly bundled payments for all post-partum care services for up to six months; and
- Outlier payments for infrequent events and unusual circumstances that result in the need for more services or more expensive services.

The bundled payments for prenatal care, labor & delivery, and post-partum care would be stratified into three risk-based categories so that higher payments are made for women who have characteristics that typically require additional or more expensive services. The woman's risk classification could change at any time, and subsequent payments would reflect the new risk category. There would be no cost-sharing for the prenatal and post-partum care services.

The Maternity Care Team would receive no payment during a month or phase of care if the Team failed to provide all evidence-based care to the woman or if a never event occurred (i.e., death of the mother, unexpected death of the infant, or iatrogenic injury to the infant). Payments to the Team would be reduced if desirable outcomes (e.g., physiologic childbirth, successful breastfeeding) were not achieved during a particular phase of care.

The APM would reduce spending and improve outcomes by enabling more women to deliver babies in birth centers rather than hospitals, reducing the frequency of Cesarean sections in low-risk births, supporting more ex-

tensive prenatal and postpartum care services for higher-risk women, and tying payments directly to outcomes.

DETAILS OF THE APM

1. Opportunities for Savings and Quality Improvement

Maternity care is one of the largest components of spending for commercial health plans and for Medicaid programs. There are a number of important opportunities for reducing unnecessary and avoidable spending on maternity care in ways that would generate savings while also improving outcomes for mothers and babies:

- Approximately one-third of babies in the United States are delivered by Cesarean section, one of the highest rates among developed countries. Payments to hospitals for C-section deliveries are significantly higher than for vaginal deliveries, so reducing the rate of C-sections would reduce spending on the delivery itself as well as reducing spending on treating complications.²⁹²
- Most vaginal deliveries in the United States take place in hospitals, even though the majority could safely take place in a birth center. Payments for vaginal deliveries in hospitals are significantly higher than for deliveries in a birth center, so increasing the proportion of births in birth centers would reduce spending and could also improve outcomes for many mothers and babies.²⁹³
- The United States has a high rate of both infant mortality and maternal mortality relative to other countries.

2. Changes in Care Delivery Needed and Associated Costs

a. New and Different Services to Be Delivered

In most large communities, birth centers exist but they are currently being underutilized. Many smaller communities, however, do not have birth centers, and a birth center would need to be created if one does not exist and if there are a sufficient number of births to sustain one.

In a growing number of small rural communities, the local hospital does not provide planned labor & delivery services, and this increases the risk of poor outcomes, particularly for higher-risk pregnancies. In these communities, the hospital would need to add the capacity for labor & delivery services. In communities where there is no hospital at all, it will be impossible to offer

hospital-based labor & delivery services in the community, and a birth center could improve outcomes and reduce the cost of deliveries in low-risk pregnancies.

In most communities, achieving better pregnancy outcomes will require approaches to delivering prenatal care and post-partum care that go beyond traditional office visits with an obstetrician or family physician, particularly for women in higher-risk categories. This would likely include proactive monitoring, education, and support to women during pregnancy and post-partum care through phone contacts and home visits. These types of services can be both effective and affordable when they are provided by healthcare professionals other than physicians, such as nurses, midwives, doulas, and community health workers.

b. Cost of Delivering New and Different Services

Although it is desirable to minimize the total number of deliveries performed in the hospital and the number of C-sections that are performed, there will always be a need for some deliveries to be performed in the hospital and for some of those deliveries to be done by C-section in order to ensure the best outcomes for the mother and baby, and it will be impossible to predict the number and timing of those deliveries. Consequently, a community will need to have at least one hospital that offers labor and delivery services and that hospital will need to incur a minimum level of fixed costs to offer 24/7 access to labor and delivery services, regardless of the actual number of deliveries that occur in the hospital. That means the average cost for hospital-based deliveries will be higher if a smaller proportion of total deliveries in the community are done in the hospital. There is a particularly significant cost for a hospital to maintain the capacity to perform surgeries such as C-sections and so the cost of each C-section will be higher if fewer C-sections and other surgeries are performed at the hospital.

Birth centers and physician/clinician practices will also need to have appropriate staffing, equipment, and facilities in order to provide maternity care services. The major driver of the costs will be personnel costs, and more personnel will be needed if there are more pregnant women in the community. A minimum level of staffing will have to be maintained in order for these services to be available when needed to address maternity care needs in the community, and that will make the average cost of these services higher in smaller communities.

There will also be startup costs involved when new services first begin. For example, if there is no birth center in the community, creating a new one will require incurring costs for construction or renovation of facilities, for recruitment and training of new staff, etc. before any services can be delivered, and there may be lower volumes of births in the birth center initially until both women and maternity care providers in the community become comfortable utilizing the birth center.

c. The Business Case for an Alternative Payment Model

An APM would be feasible if an analysis shows that the expected savings from reduced rates of Cesarean sections, hospital-based deliveries, and complications of pregnancy and childbirth would likely be larger than any increased costs associated with delivering higher-quality care during pregnancy, childbirth, and during the post-partum period. The actual costs and savings will vary from community to community depending on the number of women of childbearing age in the community and the ease of attracting and retaining high-quality maternity care providers in the community.

Table A-2a shows data for a hypothetical community with 1,000 births per year. Currently, one-third of the deliveries are by Cesarean section, and only 10% of babies are delivered in a birth center. The average payment for vaginal deliveries in the hospital is \$10,500, the average payment for a C-Section is \$12,000, the average payment for delivery in a birth center is \$4,000, and payments for prenatal and post-partum care average \$3,500. Approximately 2/3 of the costs of labor and delivery in the hospital are assumed to be fixed, and approximately 1/2 of the costs of labor and delivery in the birth center are assumed to be fixed.

As Table A-2a shows, if the rate of C-Sections decreased to 20% and the rate of deliveries in the birth center increased to 30%, the average cost of a birth in the hospital would increase because there would be fewer deliveries there, and the average cost of a birth in the birth center would decrease because of the greater number of deliveries there. (The birth center is presumed to have adequate capacity to handle the increase in births). The reduction in the number of C-Sections and the greater number of births in the lower-cost setting of the birth center would generate enough savings to preserve the operating margins of the hospital, improve the operating margin of the birth center, and still generate a net reduction in overall spending. Consequently, there would be a business case for both payers and providers to implement an Alternative Payment Model designed to support the change.

TABLE A-2a
BUSINESS CASE FOR CHANGES IN MATERNITY CARE

	Current Services, Cost, Spending			Services/Costs Under APM			
	# of Mothers	\$ Per Patient	Total	# of Mothers	\$ Per Patient	Total	Change
Services/Revenues							
Hospital Delivery							
C-Section	330	\$12,000	\$3,960,000	200	\$12,930	\$2,586,000	
Vaginal Delivery	570	\$10,500	\$5,985,000	500	\$12,930	\$6,465,000	
Birth Center							
Vaginal Delivery	100	\$4,000	\$400,000	300	\$2,700	\$810,000	
Prenatal Care	1,000	\$3,200	\$3,200,000	1,000	\$3,200	\$3,200,000	
Post-Partum Care	1,000	\$300	\$300,000	1,000	\$300	\$300,000	
Total Births/Payments	1,000	\$13,845	\$13,845,000	1000	\$13,361	\$13,361,000	-3%
Costs							
Hospital Services							
Fixed Cost			\$6,000,000			\$6,000,000	
Variable - C-Section	330	\$5,000	\$1,650,000	200	\$5,000	\$1,000,000	
Variable - Vaginal	570	\$3,500	\$1,995,000	500	\$3,500	\$1,750,000	
Margin			\$300,000			\$301,000	
Birth Center							
Fixed Cost			\$180,000			\$180,000	
Variable Cost	100	\$2,000	\$200,000	300	\$2,000	\$600,000	
Margin			\$20,000			\$30,000	
Prenatal/Post-Partum							
Cost	1000	\$3,500	\$3,500,000	1000	\$3,500	\$3,500,000	
Margin			\$0			\$0	

3. Barriers in the Current Payment System

There are a number of important barriers in current payment systems that discourage delivery of desirable maternity care services and that increase the frequency of undesirable services and poor outcomes:

- Hospitals are generally paid significantly more when Cesarean sections are performed than for vaginal delivery births, so reducing the percentage of births by C-section results in a significant reduction in hospital revenue.
- Hospitals are only paid for their labor & delivery services when a baby is actually delivered in the hospital, and the payment is the same regardless of how many births occur in the hospital, so increasing the percentage of births in birth centers could leave the hospital with insufficient revenue to cover the fixed costs of its standby capacity for hospital deliveries.
- Many health plans do not pay for births in birth centers or do not pay enough to cover their costs.
- Obstetricians are generally paid as much or more for C-sections as for vaginal deliveries, even though a normal vaginal delivery will often take significantly more time and is more likely to occur outside of normal working hours, so reducing the percentage of births by C-section will increase costs and reduce revenues to the obstetrical practice, making it more difficult to cover its costs.
- Midwives are generally paid less than obstetricians to deliver babies, so increasing the use of deliveries by midwives would reduce spending on births but could make it difficult for an obstetrician to sustain an obstetrical practice, particularly in a small community.
- Midwives need adequate payment and enough births to financially sustain a practice in the community.
- Obstetricians are generally paid a fixed “global fee” to cover all prenatal care and post-partum care services as well as delivery of the baby, so an OB practice will have more difficulty covering its costs if it encourages early and frequent prenatal care visits and if it provides additional services to high-risk women.

4. Design of the APM

The Alternative Payment Model described below could remove the barriers in the current payment system and enable and encourage maternity care providers to pursue the opportunities for improvement described above. (This is not the only way in which an APM could be designed to achieve these goals, and other approaches may be preferable for women with specific needs, or for women living in communities that are very small or have challenges in attracting and retaining maternity care providers.)

a. Defining the Eligible Patients and Physician Practices

i. Eligibility Criteria for Patients

Any pregnant woman would be eligible to receive services supported by the APM from a Maternity Care Team participating in the APM.

ii. Eligibility Criteria for Maternity Care Teams

A Maternity Care Team would be eligible to participate in the APM if it includes:

- At least one clinician (a physician, nurse practitioner, midwife, etc.) who is qualified and licensed to provide prenatal care to pregnant women;
- At least one clinician who is qualified and licensed to assist in vaginal delivery of babies;
- At least one physician who is qualified to treat women with high-risk pregnancies, to treat complications experienced during pregnancy or childbirth, and to perform Cesarean sections;
- At least one clinician who is qualified and licensed to provide care to newborn babies;
- At least one hospital that has the capability to perform Cesarean sections and treat common complications of labor and delivery; and
- At least one physician practice, hospital, clinical laboratory, or other entity that has the ability to perform any laboratory tests or imaging studies needed as part of prenatal care, labor & delivery, and post-partum care.

Ideally, the Maternity Care Team will also include at least one birth center that is licensed to provide labor and delivery services if the volume of births in the community is sufficient to support both a birth center and a hospital that offers labor and delivery services (since both will be needed to adequately serve all pregnant women).

iii. Designation of the Patient's Maternity Care Team

If a pregnant woman wanted to receive the enhanced services and accountability for cost and outcomes under the APM, the woman would designate a Maternity Care Team that is participating in the APM to provide her with high-quality prenatal care, safely deliver her baby, provide the initial care for the baby immediately following

birth, and provide the woman with high-quality post-partum care. A woman could select or change her Maternity Care Team any time prior to the beginning of labor or during the post-partum period.

Before a woman decides to designate a particular Maternity Care Team to provide her maternity care, the Team would describe the services it would commit to deliver and the approach it would use for decision-making about services in various circumstances that might arise prior to, during, and after birth. The team could also ask the woman to commit to taking actions during pregnancy and following birth that would support good outcomes (e.g., attending prenatal care exams, abstaining from smoking, drinking, and drugs during the pregnancy, contacting the Team when a problem arises, etc.). The Team could also ask the woman to only obtain maternity care services from the members of the Team unless the Team specifically recommends that the woman receive services from other providers.

In order to receive payments under the APM for a woman, the Maternity Care Team would need to verify and document that the woman was pregnant and that she had designated the Team to provide her with complete maternity care.

b. Removing the Barriers in the Current Payment System

There would be five components to the payments to the Maternity Care Team for care of the pregnant mother, covering three separate phases of maternity care delivery:

- Monthly payments for pregnancy-related services needed prior to childbirth;
- Standby capacity payments for the minimum capacity needed in hospitals to support labor and delivery services;
- Bundled/warranted payments for labor and delivery;
- Monthly payments for post-partum care services; and
- Outlier payments for infrequent events and unusual circumstances.

Each of these payments would be stratified based on factors affecting the risk of complications and conditions that typically require additional or expensive services.

Health care services for the infant beyond the immediate post-delivery care would be paid for separately because the types of care needed would not be clearly known until after birth, and different sets of providers (e.g., pediatricians) would likely be involved in the delivery of the care to the child. A separate Alternative Payment Model could be developed to address the specific opportunities and barriers to improvement for infant care.

i. Stratification of Patients Based on Need/Risk

Each participating woman would be classified by her Maternity Care Team into one of three categories: (1) Low Need/Risk, (2) Moderate Need/Risk, and (3) High Need/Risk. The criteria for this classification would be based on objective, measurable characteristics of the woman and her baby that evidence or analysis show have a significant impact on the risk of complications during birth or the types of services needed in order to ensure a good outcome. The Team would document the characteristics used to make the classification for each woman receiving maternity care from the Team.

A woman's classification could change if new risk factors or needs developed while she was receiving prenatal care, during labor and delivery, or during the postpartum care period. Payment amounts in each category would be based on the most recent classification assigned prior to the initiation of the services covered. For example, if the mother's risk level increased after the first prenatal care visit but before labor began, the Labor & Delivery Bundled Payment would be based on the higher risk category.

ii. Monthly Payments for Pregnancy Care Services

The Maternity Care Team would receive a pre-defined standard Monthly Pregnancy Care Payment each month to support delivery of all of the pregnancy-related care services a pregnant woman needed during the month. This would include prenatal care visits, laboratory tests and imaging studies, and any medical procedures. A higher amount would be paid to the Team to deliver pregnancy-related care services to women who are in the higher need/risk categories.

The Maternity Care Team would be responsible for dividing the revenues from these payments among the members of the Team to pay for the staff, equipment, travel, etc. used to deliver the prenatal care services and to pay for any services delivered by providers who are not members of the Team that deliver pregnancy-related services needed by a woman who is receiving care from the Team. The Team would have flexibility regarding the exact services it delivered, the type of personnel it used, etc., but it would be expected to follow evidence-based standards of care whenever they were applicable in order to receive payment.

The payment amounts in each category would be based on the estimated per-patient cost of pregnancy-related care services for patients in that category, with adjustments for performance as described below. (These payments would be the type of stratified, bundled, condition-based payments described in Option 6 in Section VI-A, i.e., the payments would differ based on the patient's characteristics, rather than on the services provided, and the Team receiving the payment would have the flexibility to deliver multiple services and combinations of services.)

Payment amounts would be periodically adjusted based on analyses of the actual costs incurred by Maternity Care Teams that successfully achieve the performance targets.

iii. Standby Capacity Payment for Hospital Delivery Services

The hospitals that are part of the Maternity Care Team would receive a standard, pre-defined Standby Capacity Payment for each pregnant woman who receives labor and delivery services from the Team, regardless of whether the woman delivered her baby in the hospital or not. A higher amount would be paid for women in higher need/risk categories. The revenues from these payments would be designed to support the cost of the minimum on-site and on-call staffing and equipment the hospital(s) must maintain in order to be ready to deliver a baby who needs hospital-level services for a successful delivery.

The hospitals would determine their monthly standby capacity cost for maternity care, i.e., the cost each hospital would need to incur each month to staff and equip its labor and delivery services if only one baby were delivered during the month. If a hospital uses some of the same staff and equipment for other types of patients (e.g., emergency surgery cases), then a portion of those shared costs would be assigned to the estimated cost of labor and delivery services based on the proportion of pregnant women who would be using the shared services. The monthly maternity care standby capacity cost would be divided by the total number of women delivering babies in the community served by the hospital, regardless of where the delivery actually occurred (i.e., women delivering in birth centers or at home would also be counted), to determine the per-patient standby capacity payment the hospital would receive for each woman receiving care from the Maternity Care Team.

iv. Bundled/Warranted Payment for Labor and Delivery

The Maternity Care Team would receive a single, standard, pre-defined Labor & Delivery Bundled Payment to support all of the services from any member of the Team or any other providers that the woman needed to deliver her baby and that the woman and infant needed immediately following birth. The Team would be responsible for dividing the revenues from the Labor & Delivery Bundled Payments among the individual members of the Team for the services they deliver, including facility services (at the hospital or birth center where the delivery occurred) and all professional services. The Maternity Care Team would not charge for or receive any additional payments for any services delivered to the woman or baby to treat complications of childbirth, unless the circumstances qualified for an Outlier Payment. A higher Labor & Delivery Bundled Payment amount would be paid to the Team for a woman classified in one of the higher risk categories.

The amount of the Labor & Delivery Bundled Payment for women in a particular need/risk category would be based on the expected average cost of labor and delivery services for all pregnant women in that category, except for the services or costs that would be covered by Outlier Payments and Standby Capacity Payments. The expected cost would be determined by taking the estimated cost per birth of each of the three potential modes of delivery – vaginal delivery in a birth center,

vaginal delivery in a hospital, and Cesarean section in a hospital – and weighting them by the percentage of births that could be expected to be safely delivered in a birth center in each risk category and the percentages of births that would be expected to require delivery in a hospital or a C-section. The cost per birth used for births in a birth center would be based on the estimated total cost of labor & delivery services in a birth center divided by the expected number of births in the community that could safely be delivered in a birth center. The cost per birth used for the hospital-based deliveries would be based on the hospitals' *marginal* cost of an *additional* delivery in the hospitals beyond the first birth, since the Standby Capacity Payments would already be covering a portion of the fixed costs of the hospital delivery services.

v. Monthly Payment for Post-Partum Care

The Maternity Care Team would receive a single, standard, pre-defined Monthly Post-Partum Care Payment for each month of post-partum care the woman needs, for up to six months. Higher payment amounts would be paid for women in higher need/risk categories, and higher payment amounts would be paid in the first two months than in the third through sixth months.

The Maternity Care Team would be responsible for dividing the revenues from these payments among the members of the Team to pay for the staff, equipment, travel, etc. used to deliver the post-partum care services and to pay for any services delivered by providers who are not members of the Team that deliver post-partum services needed by a woman who is receiving care from the Team. The Team would have flexibility regarding the exact services it delivered, the type of personnel it used, etc., but it would be expected to follow evidence-based standards of care whenever they were applicable in order to receive payment.

The monthly payment amounts in each risk category would be based on the estimated per-patient cost of post-partum care services for women in that category, with adjustments for performance as described below.

vi. Outlier Payments for Infrequent Events and Unusual Circumstances

The Maternity Care Team could receive one or more Outlier Payments, in addition to all other payments, for a woman who:

- experienced an unavoidable event that occurs infrequently but typically requires a significant number of additional services or additional time or costs; or
- had unusual characteristics that required additional services or additional time or costs in the delivery of typical services.

For events that occur infrequently but require predictable responses, the Team would receive a standard, pre-defined Outlier Payment. For example, Outlier Payments would be pre-defined for:

- Extended labor
- Conditions that occur during pregnancy that require a hospital admission to prevent premature labor

For unusual events, there would not be a standard pre-defined payment; instead, the amount of the Outlier Payment would be based on the actual additional costs that the Maternity Care Team incurred in delivering care to the woman during a particular phase of care (i.e., prenatal, delivery, or postnatal care). The Team would calculate the actual costs it incurred for the woman's care, and subtract the payments it had otherwise received; the Outlier Payment would then be set equal to 90% of that amount.

vii. Patient Cost-Sharing

In order to encourage women to obtain prenatal and post-partum care, there would be no copayment or other cost-sharing for the Monthly Pregnancy Care Payments, for the monthly Post-Partum Care Payments, or for Outlier Payments. Alternatively, a copayment could be charged but waived if the woman had adhered to all of the recommended actions in the plan of care developed by the Maternity Care Team.

The woman would be expected to pay a pre-defined copayment for the Labor & Delivery Bundled Payment. The copayment for the Labor & Delivery Bundled Payment would not differ based the woman's risk category. However, higher copayments would be charged in the following circumstances:

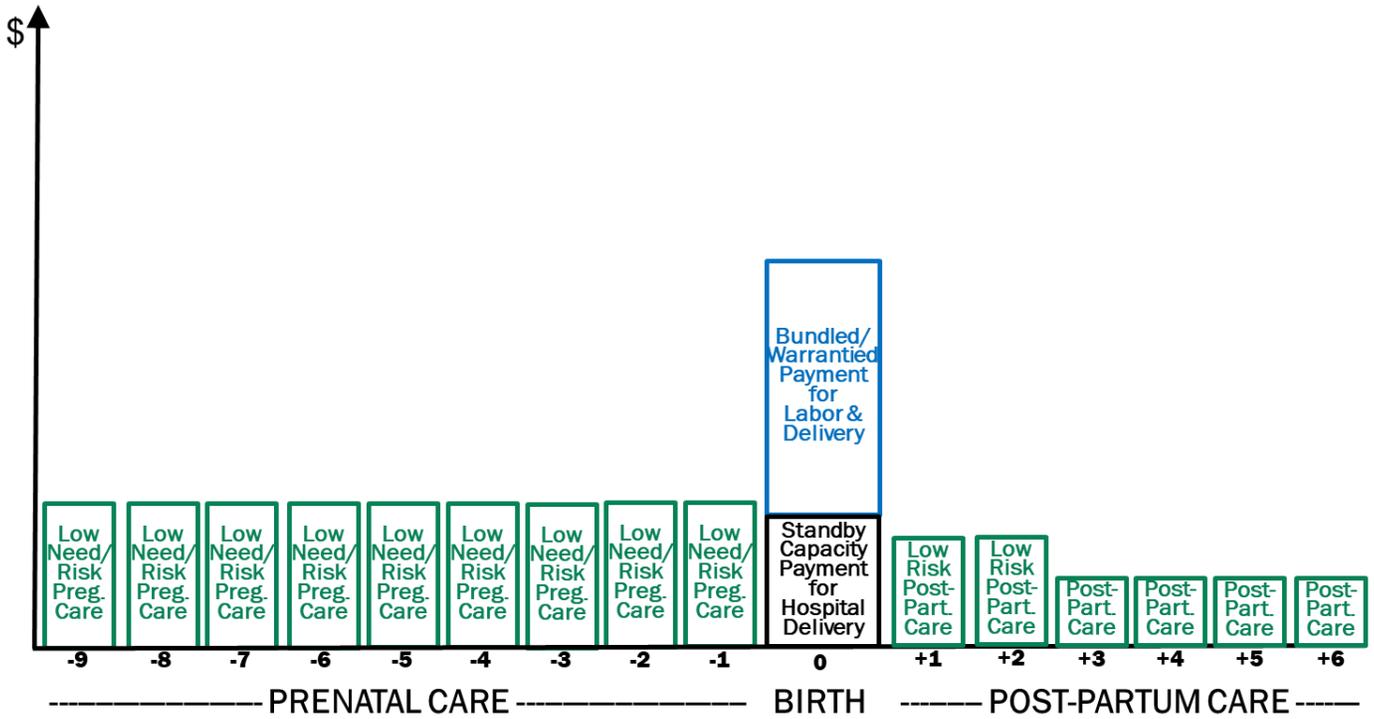
- The woman did not adhere to the prenatal care plan developed by the Maternity Care Team;
- The woman was classified in the Low or Moderate Risk categories and the Maternity Care Team felt that delivery in a birth center would be safe for the mother and baby, but the woman was only willing to give birth in a hospital;
- The woman wanted to plan a Cesarean section even though the Maternity Care Team felt that pursuing a vaginal delivery was safe or safer for the mother and baby; and/or
- The woman obtained pre-natal or post-partum care services from a provider other than the Maternity Care Team without approval from the Team.

viii. Non-Standard Charges for Maternity Care

A Maternity Care Team would be permitted to charge more than the standard payment amounts for one or more of the payment components if the Team wanted to commit to more or better outcomes than required for the standard payments. If a woman chose to use a Team that charged more, she would pay the difference between the team's charge and the standard payment amount.

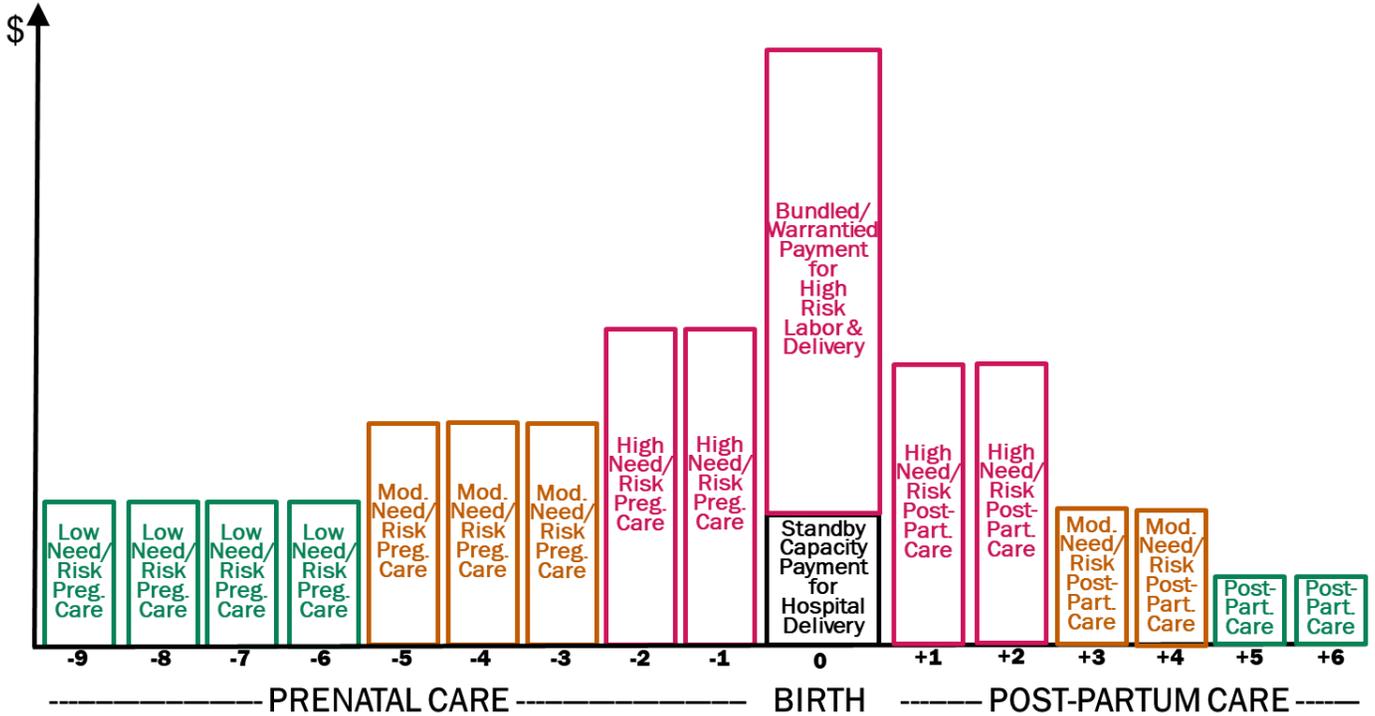
A Maternity Care Team could also charge less than the standard amount for the Labor & Delivery Bundled Payment, in which case the copayment amount would be proportionally reduced.

PAYMENT UNDER THE MATERNITY CARE APM FOR A LOW-RISK PREGNANT WOMAN



Note: Relative heights of bars are not intended to represent actual relative amounts of payment

PAYMENT UNDER THE MATERNITY CARE APM FOR A PREGNANT WOMAN WHO BECOMES HIGHER-RISK



Note: Relative heights of bars are not intended to represent actual relative amounts of payment

c. Creating Accountability for Utilization and Spending

In return for receiving the payments specified in the previous section, the Maternity Care Team would be responsible for delivering (or paying other providers to deliver) all maternity care services the woman needed during the phase of care for which the Team was receiving payment under the APM. No other payments would be made to the Team for maternity-care related services delivered to the woman. If a provider other than a member of the Team delivered a maternity care-related service to a woman who had designated the Team for her care and that other provider billed the woman's health insurance plan for the service, the plan would pay for that service and deduct the amount it paid from the payments due to the Maternity Care Team for the phase of care in which that service was delivered.

d. Creating Accountability for Quality

i. Measures of Quality/Outcomes

There are many dimensions to the quality of maternity care. Good measures are not currently available for all aspects of maternity care quality, the relative importance of measures differs for different women, and the knowledge of how to reliably achieve them also varies. To address this, three groups of quality measures would be used:

- **Never Events.** These are serious events that would be always viewed as unacceptable. The patient-level measure would be whether the never event occurred, and the population-level measure would be the proportion of patients experiencing the never event. The never events would include:
 - ◆ Death of the mother
 - ◆ Death of the infant
 - ◆ Iatrogenic injury to the infant
- **Desirable Outcomes.** These are outcomes that are desirable to achieve and that are believed to be feasible to achieve for the individual patient. The patient-level measure would be whether the desirable outcome occurred, and the population-level measure would be the proportion of patients achieving the outcome, excluding those patients for whom the outcome was impossible or contraindicated. These would include:
 - ◆ Physiologic ("natural") childbirth
 - ◆ Successful breastfeeding
 - ◆ Lack of postpartum depression
- **Evidence-Based Care.** These would be services for which there is evidence that delivery of the service increases the likelihood of desirable outcomes by a large amount. The patient-level measure would be whether the patient received care that followed evidence-based care guidelines. If the guidelines indicated that a service was not necessary or appropriate for all patients, the measure would only apply to patients with the characteristics for whom the service was appropriate.

ii. Target for Quality Measures

- Never Events: The Target would be a rate of 0%.
- Evidence-Based Care: The Target would be 100%.
- Desirable Outcomes: The Target would be 100%.

iii. Mechanism for Accountability

Never Events

If a never event occurred, the Maternity Care Team would not receive any payments at all and would refund any payments that had already been made prior to the occurrence of the never event.

Evidence-Based Care

If the services delivered by the Maternity Care Team during a particular month or phase of care failed to meet evidence-based guidelines, the Team would not receive payment for that month or phase. For example, if the Team failed to perform a prenatal test required by evidence-based guidelines, the Team would not receive the Monthly Pregnancy Care Payment for the month in which the test should have been performed.

Desirable Outcomes

The payment for a particular month or phase of care would be reduced if one or more of the specified desirable outcomes for that month or phase of care was not achieved. The amount of the reduction associated with each outcome would be specified in advance, and the amounts would differ for different outcomes. If there are multiple desirable outcomes to be achieved in a particular month or phase, there would be a cap on the total amount by which the payment could be reduced if none of the outcomes were achieved.

e. Ensuring the APM Design Supports the Business Case

Table A-2b shows examples of payment amounts that would support the business case. The payment for labor and delivery would be based on the need/risk characteristics of the mother and baby, not where the delivery occurred or the method of delivery used. The hospital would receive a standby capacity payment for each delivery, including for deliveries in the birth center, in order to help support its fixed costs. The total payment for delivery for a low-risk mother (\$3,500 + the \$1,200 standby capacity payment) would be similar to what a woman would have paid for a birth center delivery previously, but without the risk of a much larger payment if the woman had to be transferred to the hospital during the delivery. The total payment for a medium-risk mother (\$9,000 + \$1,200) would be similar to what is currently being paid for a vaginal delivery in the hospital. The total payment for delivery of a baby by a high-risk woman would be much higher than it is today, more accurately reflecting the higher cost. Total spending (based on the expected number of births and distribution of risk categories in the community) would be lower than it is currently.

TABLE A-2b
HOW THE APM WILL ACHIEVE THE BUSINESS CASE FOR IMPROVED MATERNITY CARE

	SERVICES/SPENDING UNDER CURRENT FFS			SERVICES/SPENDING UNDER ALTERNATIVE PAYMENT MODEL					
	#	\$ Per Patient	Total	#	\$ Per Month	# Mo.	\$ Per Patient	Total	
Services/Revenues									
Hospital Delivery				Delivery					
C-Section	330	\$12,000	\$3,960,000	High-Need	200		\$15,500	\$3,100,000	
Vaginal Delivery	570	\$10,500	\$5,985,000	Med. Need	500		\$9,000	\$4,500,000	
Birth Center				Low Need	300		\$3,500	\$1,050,000	
Vaginal Delivery	100	\$4,000	\$400,000	Standby	1,000		\$1,200	\$1,200,000	
Prenatal Care	1,000	\$3,200	\$3,200,000	Prenatal Care					
				High-Need	200	\$422.22	9	\$3,800	\$760,000
				Med. Need	500	\$355.56	9	\$3,200	\$1,600,000
				Low Need	300	\$311.11	9	\$2,800	\$840,000
Post-Partum Care	1,000	\$300	\$300,000	Post-Partum Care					
				High-Need	200	\$66.67	6	\$400	\$80,000
				Med. Need	500	\$50.00	6	\$300	\$150,000
				Low Need	300	\$41.67	6	\$250	\$75,000
Total Spending	1,000	\$13,845	\$13,845,000	Total Spending	1,000			\$13,355	\$13,355,000
Costs									
Hospital Services				Hospital Services					
Fixed Cost			\$6,000,000	Fixed Cost				\$6,000,000	
Variable Cost C-Section	330	\$5,000	\$1,650,000	Variable Cost C-Section	200		\$5,000	\$1,000,000	
Variable Cost Vaginal	570	\$3,500	\$1,995,000	Variable Cost Vaginal	500		\$3,500	\$1,750,000	
Birth Center				Birth Center					
Fixed Cost			\$180,000	Fixed Cost				\$180,000	
Variable Cost	100	\$2,000	\$200,000	Variable Cost	300		\$2,000	\$600,000	
Prenatal/Post-Partum				Prenatal/Post-Partum					
Cost	1,000	\$3,500	\$3,500,000	Cost	1,000		\$3,500	\$3,500,000	
TOTAL COST			\$13,525,000	TOTAL COST				\$13,030,000	
Margin			\$320,000	Margin				\$325,000	

5. Operationalizing the APM

In order for a Maternity Care Team to be paid for delivering maternity care services under the APM, the team would submit claims forms for each eligible woman using a series of new billing codes. Penalties for failure to achieve Desirable Outcomes would be operationalized as “withholds,” i.e., the default amount of payment for a service code would be determined by calculating the estimated cost of delivering services and subtracting the maximum penalty for failure to achieve the outcomes. Additional codes would be created to enable the Team to recoup the penalty when one or more Desirable Outcomes were actually achieved.

Monthly Pregnancy Care Payments

- MC011: one month of pregnancy care to a woman who meets the criteria for the Low Need/Risk category;
- MC012: one month of pregnancy care to a woman who meets the criteria for the Moderate Need/Risk category; and
- MC013: one month of pregnancy care to a woman who meets the criteria for the High Need/Risk category
- MC014: occurrence of an infrequent and unavoidable condition that requires additional time or services during pregnancy
- MC015: occurrence of an infrequent and unavoidable condition that requires inpatient care during pregnancy
- MC016: occurrence of unusual circumstances requiring additional services or costs
- MC101-MC108: achieving specified Desirable Outcomes during a month (each outcome would be assigned a separate code)
- MC109: maximum additional payment for achieving Desirable Outcomes. If the Maternity Care Team had achieved multiple Desirable Outcomes, it would submit individual codes (MC101-MC108) for each of those outcomes, and if the total additional payments for those codes exceeded the maximum additional payment per patient, the Team would also submit code MC109 and the payment would be made for that code instead of the others. (All of the codes would still be submitted so it was clear which outcomes had been achieved.)

Bundled/Warranted Payments for Labor and Delivery

- MC021: labor & delivery services to a woman who meets the criteria for the Low Need/Risk category;
- MC022: labor & delivery services to a woman who meets the criteria for the Moderate Need/Risk category; and
- MC023: labor & delivery services to a woman who meets the criteria for the High Need/Risk category
- MC024: extended labor

- MC025: occurrence of an infrequently-occurring condition that requires additional time or services during labor and delivery
- MC026: occurrence of unusual circumstances requiring additional services or costs during labor and delivery
- MC201-MC208: achievement of specified Desirable Outcomes during labor & delivery
- MC209: maximum additional payment for achieving Desirable Outcomes.

Monthly Post-Partum Care Payments

- MC031: one month of post-partum care during the first two months following delivery for a woman who meets the criteria for the Low Need/Risk category
- MC032: one month of post-partum care during the first two months following delivery for a woman who meets the criteria for the Moderate Need/Risk category
- MC033: one month of post-partum care during the first two months following delivery for a woman who meets the criteria for the High Need/Risk category
- MC034: one month of post-partum care during the third through sixth months following delivery for a woman who meets the criteria for the Low Need/Risk category
- MC035: one month of post-partum care during the third through sixth months following delivery for a woman who meets the criteria for the Moderate Need/Risk category
- MC036: one month of post-partum care during the third through sixth months following delivery for a woman who meets the criteria for the High Need/Risk category
- MC037: occurrence of an infrequent, unavoidable condition that requires additional time or services during the post-partum care period
- MC038: occurrence of an infrequent, unavoidable condition that requires extended inpatient care or a readmission following delivery
- MC039: occurrence of unusual circumstances requiring additional services or costs
- MC302-MC308: achievement of specified Desirable Outcomes during post-partum care
- MC309: maximum additional payment for achieving Desirable Outcomes during post-partum care

Submission of Claims

The date of service on the claim would be the last day of the month in which the pregnancy care or post-partum care services were delivered (for the Pregnancy and Post-Partum Care Payments), the day on which the baby was delivered (for the Labor & Delivery Payments), or the day on which the Desirable Outcome was achieved or documented (for the bonus payments based on achieving Desirable Outcomes).

Submission of a claim form with one of these billing codes would represent a certification by the Maternity Care Team that:

- The woman met the eligibility criteria for the APM and for the assigned Need/Risk category.
- The team had delivered services to the woman that met all required evidence-based standards for that phase and month of care.
- The mother and the baby had both survived without any iatrogenic injuries.

If the Maternity Care Team wished to charge women more than the amount that would be paid by their health plans, the Team would publish its charge for each of the billing codes, and the woman would agree to those charges at the time that she was enrolling to receive maternity care from the Team. A Team that charged a higher amount would charge the same amount to all women, regardless of their health insurance plan, and the Team would bill the woman for the difference between the charge and the amount paid by her health insurance plan.

On a quarterly basis, the Maternity Care Team would calculate its performance on all of the quality measures (Never Events, Evidence-Based Care, and Desirable Outcomes). These rates would be calculated separately for women in each of the three need/risk categories. The rates would be provided to the Team's patients and to the health insurance plans for those patients.

The Maternity Care Team would make information about its performance on the quality measures and its charges for services publicly available so that women who were seeking a Maternity Care Team could compare the cost and performance of different Teams.

Standby Capacity Payments

Because the hospitals participating in the APM would receive a Standby Capacity Payment for all women delivering babies as part of the APM regardless of whether the baby was actually delivered in the hospital, it would be difficult for the hospital to bill directly for all of these payments. Instead, since the payments would be made if and only if a Maternity Care Team received a Labor & Delivery Bundled Payment, the submission of a claim by a Maternity Care Team to a participating health insurance plan for one of those payments would also automatically trigger a Standby Capacity Payment from the health insurance plan to each participating hospital.

To distinguish the payment made to the Maternity Care Team from the Standby Capacity Payment made to a hospital, a modifier would be added to the codes listed earlier:

- -OP: Labor & Delivery Bundled Payment to the Maternity Care Team
- -IP: Standby Capacity Payment to a hospital

For example, if a Maternity Care Team submits a claim with a MC021 code for a birth in a birth center, the health plan would issue a payment to the Maternity Care Team with the amount assigned to the MC021-OP code and modifier, and the health insurance plan would also

issue a payment to each participating hospital with the amount assigned to the MC021-IP code and modifier.

6. Implementing the APM

a. Obtaining Participation by Payers, Providers and Patients

The Maternity Care APM would have a number of advantages that should encourage payers to implement the APM, encourage providers to participate in the APM, and encourage pregnant women to seek care from providers who are participating in the APM.

i. Advantages for Payers

- Participating health insurance plans could reduce spending on plan members who are pregnant or have recently delivered a baby without negative impacts on their members.
- Health insurance plans could implement the APM with minimal administrative costs by creating new billing codes in the payer's existing claims payment system.

ii. Advantages for Maternity Care Providers

- Maternity care providers would have the flexibility to deliver services to their patients in the ways that are most feasible for the providers and most effective for their patients, including delivery of appropriate services by obstetricians, midwives, nurses, doulas, and/or community health workers.
- Participating maternity care providers would receive higher payments to cover the additional time they would spend with women who begin prenatal care early in their pregnancies, who have higher risk pregnancies, and/or who need additional time to deliver their baby through natural childbirth. Obstetricians could focus their time more on high-risk pregnancies and still receive enough revenues to sustain their practices.
- Participating maternity care providers would be responsible for following evidence-based clinical guidelines and for avoiding never events, but they would not be penalized for delivering care that their patients needed nor would they be penalized for increases in the amounts that other providers charged for their services or for increases in the prices of drugs and medical devices.
- Participating physician and midwife practices would know when to expect payment and how much to expect based on the bills they submit to payers and the cost-sharing charged to patients. The largest financial loss a practice could experience would be the loss of the payments under the APM.
- Participating hospitals would no longer have all of their revenues tied to the number of babies delivered in the hospital and the method of delivery; the hospital could support efforts to reduce inappropriately high C-section rates and to encourage more births in birth centers without losing money by doing so.

-
- Participating maternity care providers could bill for services using their standard billing systems.

iii. Advantages for Pregnant Women

- Women would have the choice of whether to receive maternity care services supported by the APM based on a clear understanding of the services they would receive, the actions they would need to take, and the results they could expect to achieve.
- Women would have the choice about where to deliver their babies and the method of delivery that would be used.
- Women could change maternity care providers before and after delivery if they wished to do so.
- Women would know that their maternity care providers would be rewarded for achieving good outcomes but would have no financial incentive to withhold needed care.
- Women would know how much they would need to pay for maternity care services before choosing to receive them.

b. Finalizing the APM Parameters

A “beta test” of the APM will likely be needed with willing providers in order to finalize several key parameters of the APM:

- **Criteria defining the categories of need/risk.** The categories should be defined so that they distinguish which women will be at higher risk of complications and which women will need more time and care management services from maternity care providers in order for them to follow evidence-based care guidelines and to improve patient outcomes. However, data may not be available on all of the factors that would be expected to affect need and risk, and the APM will need to be implemented in order to enable those data to be collected.
- **Dollar amounts of the various payments.** The payment amounts in each phase of care and for each need/risk level should be based on the cost of the services that would be delivered to women in that phase and level, but the cost of the services will depend on the number of patients the providers can manage and the number of patients in each of the need/risk categories, and this can only be estimated after the services are actually implemented with support from the APM. For example, the cost of births in birth centers might be reduced below current levels if a larger number of women begin to use them.
- **Benchmark rates of desirable outcomes.** Data are not currently being collected for many types of desirable outcomes for maternity care because there is no means of paying for the costs of doing so. Consequently, performance targets and payment amounts for many types of desirable outcomes can only be determined after services under the APM begin.

Best estimates of these parameters would be used to initiate the beta test process, and the participants would gather and share data from their actual experience in implementing care changes with payments under the APM in order to make adjustments to the parameters.

APM #3:

Management of a Chronic Condition

Every chronic disease is different. Different treatments are needed for different diseases, the cost and effectiveness of treatments varies across different diseases, there are more alternative treatments for some diseases than others, and the severity of complications from overtreatment and undertreatment vary. There are additional differences and complexities when patients have additional health problems or face barriers in accessing healthcare services.

However, despite these differences, there are also many similarities in the opportunities for improvement, in the barriers that current payment systems create to improving care delivery, and in the ways in which payments could be changed to support higher-quality, more affordable care across a wide range of chronic diseases and combinations of diseases. This section will focus on some of the opportunities, barriers, and payment changes that are common to a number of different chronic diseases and combinations of diseases and how an Alternative Payment Model might address them. For simplicity, the term “chronic condition” will be used here to describe either a single chronic disease or a combination of two or more chronic diseases that need to be managed in close coordination.

OVERVIEW OF THE APM

Under this APM, an individual who has the symptoms of a serious chronic disease or who has been diagnosed with the disease would choose one or more teams of providers that are participating in the APM to diagnose, treat, and manage the individual’s condition. Seven types of payments would be available under the APM in order to match the different kinds of services that the patient would need and the different outcomes that can be achieved during five different phases of care:

- 1. Diagnosis and Initial Treatment.** A Diagnosis Team would receive a one-time bundled Diagnosis and Initial Treatment Payment to cover most of the services needed to determine if the patient has the chronic disease, and if so, to treat the disease for an initial period of time. The payment would be higher for those patients who are diagnosed with the disease and initiate treatment.
- 2. Continued Treatment for Patients with Well-Controlled Conditions.** A Treatment Team would receive a quarterly bundled Treatment and Care Management Payment to provide appropriate services for patients whose condition can be well-controlled with standard medications or other treatments. In some cases, the Treatment Team would be the same as the Diagnosis Team and in other cases it might be a different group of providers.
- 3. Continued Treatment for Patients With Difficult-to-Control Conditions.** If the patient’s condition proved

difficult to control during the initial treatment period or if it could only be controlled using special medications or treatments that require careful monitoring, a Treatment Team would receive a quarterly bundled Treatment and Care Management Payment to provide appropriate services. The payment amounts would be higher than for patients with well-controlled conditions, reflecting the greater risk of complications and higher level of services needed.

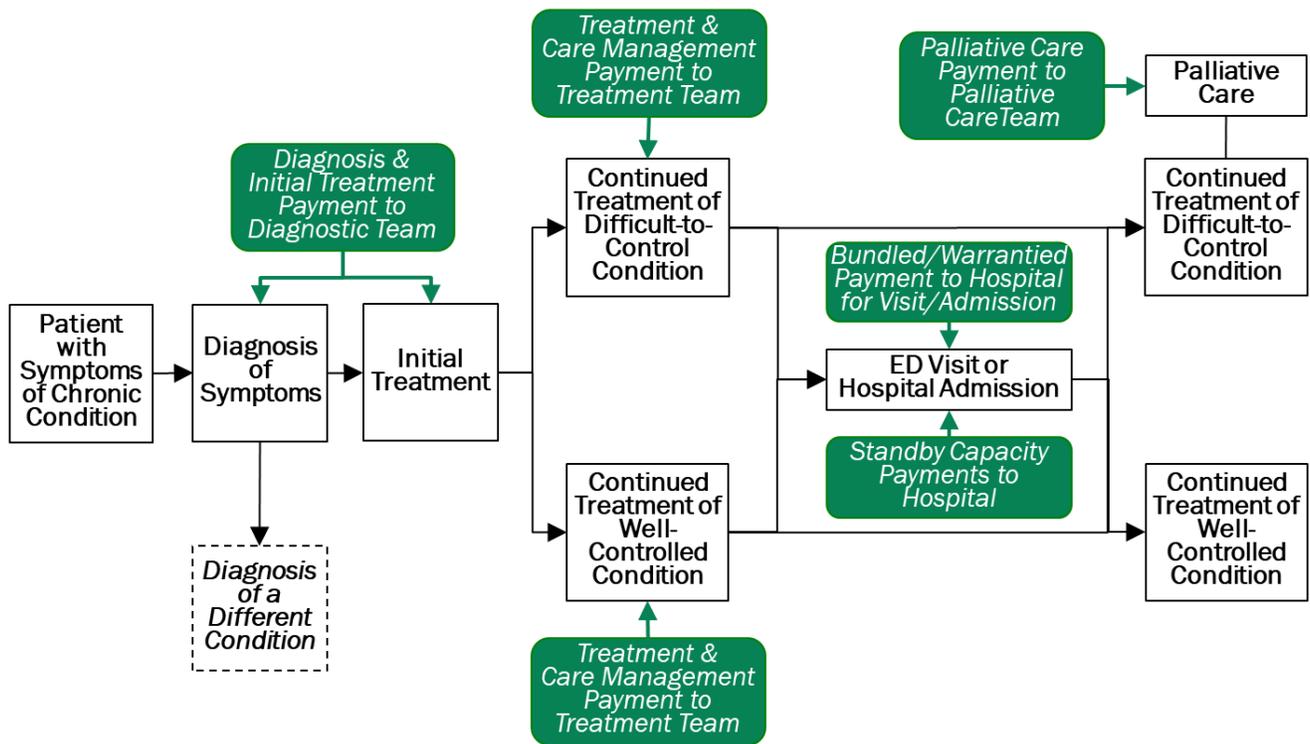
- 4. Hospitalization for an Exacerbation of the Condition.** Hospitals would receive three separate types of payments to cover the costs of their services to patients who need to be hospitalized for exacerbations of their condition:
 - a. A Standby Capacity Payment for each patient who has the chronic condition, regardless of whether they needed to be hospitalized.
 - b. A Bundled/Warranted Payment if the patient requires a visit to the Emergency Department or an inpatient admission for symptoms related to their chronic condition. This would cover all of the costs of the ED visit or hospital admission and any post-acute care services needed for 30 days following discharge that were not provided by the patient’s Treatment Team.
 - c. An Outlier Payment if a patient required an unusually large number of services.
- 5. Palliative Care for an Advanced Condition.** For patients whose condition has reached an advanced stage, a Palliative Care Team could receive a monthly Palliative Care Payment to provide palliative care services to the patient in addition to any treatment or care management services the patient was receiving from a Treatment Team.

The payments in each phase would be stratified into several need/risk-based categories so that higher payments are made for patients who have characteristics that typically require additional or more expensive services. The patient’s need/risk classification could change at any time, and subsequent payments would reflect the new need/risk category.

Diagnosis Teams, Treatment Teams, hospitals, and Palliative Care Teams would receive no payment for a patient if the Team failed to meet evidence-based care standards in providing services to that patient. Payments to a Team or hospital would be reduced if desirable outcomes were not achieved. Treatment Teams would receive no payment for low- and moderate-risk patients if the patient visited the ED or was hospitalized.

The APM would reduce spending and improve outcomes by reducing the rate of avoidable emergency department visits and hospital admissions and by reducing the utilization of unnecessary medications, tests, and other services.

PHASES OF CARE AND ASSOCIATED PAYMENT COMPONENTS IN AN APM FOR MANAGEMENT OF A CHRONIC CONDITION



DETAILS OF THE APM

1. Opportunities for Savings and Quality Improvement

Treatments for chronic diseases represent a large proportion of total healthcare spending for most payers, particularly Medicare. There are a number of important opportunities for reducing unnecessary and avoidable spending on patients who have chronic diseases in ways that would improve outcomes for the patients:

- Many individuals visit multiple physicians and undergo repeated or unnecessary testing before receiving a diagnosis and initiating treatment.
- Many individuals are incorrectly diagnosed, resulting in unnecessary, expensive, and potentially harmful treatment for the wrong disease and delays or failure to receive the correct treatment.
- A variety of expensive new drugs have been developed to treat chronic diseases; these drugs are more effective than traditional drugs for some patients, but using them for every patient increases spending without any benefit for many patients.
- Many patients with a chronic illness are admitted to the hospital because the symptoms of their illness become uncontrolled and sufficiently severe that they require inpatient treatment. Reducing the frequency of these unplanned, expensive hospital admissions would reduce spending for both the patient and their health insurance plan. In addition, avoiding the hospitalizations will reduce the risk of the patient developing additional health problems during their hospital stay (e.g., a hospital-acquired infection) that could require additional treatment and spending.
- Many patients who are hospitalized for a chronic disease exacerbation, particularly older patients, spend time in a skilled nursing facility (SNF) after discharge rather than returning directly home. These SNF stays are also expensive, and they can also cause additional health problems, so finding ways to provide post-acute care services in the home can be better for patients as well as reducing spending for payers.
- In some cases, patients are not receiving treatments or assistance that could slow the progression of their disease and delay the need for more intensive and expensive treatments.
- Patients with advanced illnesses often receive expensive treatments that have little clinical benefit and can result in reduced quality of life and increased rates of hospitalization in the days and months prior to their death.

2. Changes in Care Delivery Needed and Associated Costs

a. New and Different Services to Be Delivered

A variety of demonstration projects have shown that delivering additional services and delivering services in different ways can improve care and reduce spending in these various opportunity areas. For example:

- Electronic consults and telehealth visits with specialists can enable many patients to be diagnosed more efficiently and accurately, particularly patients who live in rural areas where there are shortages of specialists and patients who have difficulty traveling to medical appointments.
- Taking the time to consult evidence-based guidelines and to engage in shared-decision making processes with patients can enable physicians to reduce utilization of expensive treatments that would have little or no benefit for the patient.
- Patient education and self-management supports can help patients reduce the frequency and severity of exacerbations.
- Proactive monitoring of patient symptoms and rapid response to exacerbations by physician practices can reduce the severity of problems and the need for emergency department visits and hospital admissions.
- Delivery of home-based services can avoid the need for hospitalizations and skilled nursing facility stays.
- Palliative care services can help patients with advanced illnesses control the severity of symptoms and reduce the need for expensive treatments.

b. Cost of Delivering the New Services

The exact costs of delivering new and different services will vary from community to community and provider to provider depending on the type of staff used to deliver services, the number of patients with the chronic conditions, the population density of the community, and other factors. For example, home-based services are more expensive to deliver in rural areas because of the long distances between homes, the greater difficulties of attracting staff with specialized skills, and the limited access to public transportation and broadband internet services.

The cost of services will be lower if they can be used for a larger number of patients, so the more types of chronic diseases that can enable patients to qualify to participate, the lower the cost of the services can be, particularly if there are ways for multiple providers to share the same staff to deliver services.

There will also be startup costs involved when new services first begin. New staff will need to be recruited and trained before they can deliver any services, and initial caseloads may be lower while patients are first enrolling in the service.

Although it will be desirable to minimize the number of patients who are hospitalized, there will always be a

need for some patients to receive inpatient care or emergency medical care on short notice, and the fixed costs associated with maintaining that capacity will cause the average cost of the inpatient and ED services to increase when the rate of utilization decreases.

c. The Business Case for an Alternative Payment Model

An APM will be feasible for a particular chronic disease if an analysis shows that the expected savings from reduced spending on office visits, tests, medications, procedures, emergency department visits, hospital admissions, skilled nursing facility stays, etc. would be larger than the cost of delivering the new and different services needed to achieve those savings.

3. Barriers in the Current Payment System

In general, there is either no payment at all for the kinds of new and different services discussed above, or the payments that are available are insufficient to cover the costs of delivering those services in various circumstances. For example:

- Physician practices are paid for face-to-face visits between a clinician and the patient, but they are generally not paid for assistance delivered through a phone call or email. There is generally no payment for services delivered to a patient by a nurse, educator, or community health worker unless it is under the direct supervision of a physician or other clinician. If a physician practice can address a patient's need without the patient making an office visit with a clinician, revenues to the practice will decrease even though costs will not change.
- There is generally no payment to support telephone and electronic consultations between physicians to discuss and resolve alternative diagnoses, to determine what to do when standard treatments are ineffective, and to coordinate treatment plans and services for patients with multiple chronic conditions or multiple health problems. Specialists are only paid for consultations when patients visit the specialty practice, and primary care practices are only paid when patients visit the PCP, so if the patient only visits one of the practices, the other will lose revenues.
- Hospitals are not paid for maintaining the minimum capacity needed to treat patients in the emergency department and in an inpatient unit; the hospital only receives revenue to cover those costs when a patient actually visits the ED or is admitted to the hospital, and the payment is the same regardless of how many patients visit the ED or are admitted. As a result, reducing the frequency of ED visits and hospital admissions could leave the hospital with insufficient revenue to cover the fixed costs of its standby capacity.
- There is generally limited or no payment for various kinds of intensive home-based services that could serve as an alternative to an admission to a hospital or a skilled nursing facility.

- There is generally no payment to support delivery of community-based palliative care services in conjunction with treatment; although hospice programs deliver palliative care services, patients are generally required to forego treatment in order to be eligible for hospice services.

4. Design of the APM

For many chronic conditions, care delivery can be divided into five phases:

- 1. Diagnosis and initial treatment.** The first phase of care is focused on assessing symptoms to determine whether the patient has the chronic condition, establishing an initial treatment plan if they do, and delivering the initial treatment. In some cases, it is difficult to ensure an accurate diagnosis other than by determining whether treatment is effective (e.g., if there is no test that can definitively establish that the patient has the disease or if there is no test that is safe, feasible, and affordable to administer routinely), so diagnosis and initial treatment will often need to be considered as a single phase.
- 2. Continued treatment for patients with a well-controlled condition.** Ideally, after an effective treatment is identified during the initial phase of care, the patient's chronic condition will be able to be well-controlled through continued use of that treatment and through basic care management services.
- 3. Continued treatment for patients with a difficult-to-control condition.** Some patients may not respond well to standard, low-risk treatments and they may require special treatments that have higher risks of complications, or they may need more intensive care management services or services from additional or different providers in order to adequately address symptoms and minimize exacerbations of their condition.
- 4. Hospitalization for an exacerbation of the condition.** Although the goal of chronic condition treatment and management would be to avoid hospitalizations, it is likely that at least some patients will need to be hospitalized for exacerbations of their condition, and when that occurs, they will need to receive quality inpatient care at the most affordable cost.
- 5. Palliative care for an advanced condition in addition to or instead of treatment.** Patients with more advanced disease will likely experience more severe symptoms that cannot be adequately controlled through standard treatments, and in addition to treatment, they will need palliative care, i.e., services to address their symptoms. For chronic diseases that normally progress to death, patients will need to have effective end-of-life care when treatment is no longer effective and/or has unacceptable side effects.

Different services will need to be delivered by different providers during each of those phases. The costs and outcomes in each phase will differ, and so the structure and amount of payments will also need to be different in each phase.

The payments described below for each phase should be viewed as a general template for an APM that could be used to support high-quality care for many different types of chronic diseases and combinations of disease. Additions and modifications would likely be needed in order to fully address all of the opportunities for improvement associated with a specific chronic condition and to align care delivery and payment with the unique characteristics of patients, treatments, and outcomes associated with that condition. However, building APMs for different chronic conditions from a common template will make it easier for payers and multi-specialty providers to implement the APMs and will also make it easier to structure services for patients with multiple diseases.

a. Diagnosis and Initial Treatment

i. Eligibility of Patients and Designation of Diagnostic Team

Patients would be eligible to receive services supported by the APM in this phase of care if they have not been diagnosed with the particular chronic condition that is the focus of the APM but if they are experiencing symptoms that could be due to that chronic condition.

A patient who is experiencing the symptoms would choose a Diagnostic Team that participates in the APM to determine whether the patient has the chronic condition and to provide initial treatment if they do. Diagnostic Teams could vary in their willingness and ability to (1) diagnose all potential causes of symptoms or merely to determine whether symptoms are due to one of a specified group of conditions and (2) provide initial treatment for the condition that is diagnosed themselves or refer the patient to other providers for the initial treatment.

Before a patient chooses a Diagnostic Team to provide services, the Team would describe the services that it would deliver and the standards for service delivery that it committed to meet. The Team could also ask the patient to commit to actions that would support efficient and accurate diagnosis and good outcomes from initial treatment. In particular, the Team could ask the patient to only obtain diagnostic and treatment services related to their symptoms or condition from the members of the Team unless the Team specifically recommends that the patient receive services from other providers.

ii. Payments to the Diagnostic Team

The Diagnostic Team would receive a one-time bundled Diagnosis and Initial Treatment Payment to support all of the services needed to determine whether the patient has the particular chronic condition (or one of a group of chronic conditions) that is the focus of the APM and to provide initial treatment services if the condition is present.

The payment would be expected to cover the costs of office visits, laboratory tests, imaging studies, etc. used for diagnosis. If the patient is diagnosed with the condition, a higher payment would be made to cover the

costs of office visits or other patient contacts, tests and imaging studies, and procedures used to treat the patient during an initial period of time. In addition, the payment for initial treatment would be higher for patients with more severe symptoms who need more frequent treatment.

Physician practices on the Diagnostic Team would not bill or be paid for office visits or other traditional Evaluation & Management services; revenues would come through the bundled payment. If the patient receives diagnostic or treatment services related to the symptoms or condition from providers other than the members of the Diagnostic Team during the period of time in which the Diagnosis and Initial Treatment Payment is in effect, all or part of the payments the payer makes to those providers would be deducted from the Diagnosis and Initial Treatment Payment.

In general, the costs of any medications prescribed for treatment would not be included in the bundled payment, but would be paid for separately by the patient or the patient's insurance plan. Laboratory tests or imaging studies that are very expensive and only needed in certain circumstances would also be paid for separately. If the medications are purchased and administered by the Diagnostic Team, or if the expensive tests/studies are performed by the Diagnostic Team, the separate payment would be designed to cover the out-of-pocket costs incurred by the Team (e.g., the acquisition cost of the medication or of the materials required for tests).

The length of the initial treatment period would be based on the expected amount of time required to determine whether treatment is effective or which of several alternative treatments is most effective.

The Diagnostic Team would be responsible for dividing the Diagnosis and Initial Treatment Payment among the Team members to cover the costs they incur in delivering specific types of services to the patient. For example, a Diagnostic Team might consist of a primary care practice and a specialty physician practice located in a distant city; the specialty physician practice might take responsibility for determining the diagnosis and developing the treatment plan, and the primary care practice would supervise the initial treatment of the patient, but the primary care practice would consult with the specialty practice if the initial treatment is not working in order to determine how the treatment plan should be changed. One option would be for the specialty physician practice to bill the patient's health insurance plan for the Diagnosis and Initial Treatment Payment, and then use a portion of that Payment to pay the primary care practice for supervising the treatment of the patient. Another option would be for a primary care practice to bill for the payment, but contract with a specialty practice to assist in the diagnosis and treatment planning process.

iii. Accountability for Utilization and Spending

The Diagnostic Team would be held accountable for utilization and spending in two ways:

- **Bundled Payment:** The structure of the bundled payment would make the Diagnostic Team directly accountable for utilization and spending on all planned services related to diagnosis and treatment other than the out-of-pocket costs of medications and infrequent, expensive tests.
- **Evidence-Based Care:** The Diagnostic Team would be required to follow evidence-based clinical guidelines in determining which tests, medications, and procedures to deliver or order. If the Team failed to follow the guidelines for a patient and did not document the reason for deviating from the guidelines, it would not receive the Diagnosis and Initial Treatment Payment for that patient.

iv. Accountability for Quality and Outcomes

The Diagnostic Team would be held accountable for quality and outcomes in two ways:

- **Evidence-Based Care Standards:** In addition to defining which medications and tests were appropriate, the evidence-based clinical standards or guidelines would also define any other services or methods of delivery of services that had been demonstrated to result in more accurate diagnosis or better treatment outcomes for patients. If the Team failed to follow the guidelines for a patient and did not document the reason for deviating from the guidelines, or if the Team failed to meet the service standards that it had committed to meet when the patient chose the Team to deliver care, the Team would not receive the Diagnosis and Initial Treatment Payment for that patient.
- **Desirable Outcomes:** One or more measures of successful treatment would be defined that are relevant to the specific chronic condition being treated. The Diagnosis and Initial Treatment Payment would be reduced by a pre-defined amount for an individual patient when a desirable outcome did not occur or when an undesirable outcome did occur.

v. Patient Cost-Sharing

The patient would be responsible for paying a fixed copayment for the services supported by the Diagnosis and Initial Treatment Payment that are delivered by the Diagnostic Team or by providers approved by the Diagnostic Team. This copayment would be set at a level that is at or below the total of the cost-sharing amounts that the patient might expect to pay currently for individual services they would receive as part of the diagnosis and initial treatment phase of care.

If the patient receives diagnostic or treatment services from other providers without approval from the Diagnostic Team during the period of time that the Diagnosis and Initial Treatment Payment is in effect, the patient would pay additional cost-sharing for those services.

b. Continued Treatment for Patients with a Well-Controlled Condition

i. Eligibility of Patients and Designation of Treatment Team

Patients would be eligible to receive services supported by the APM in this phase of care if they have been diagnosed with one of the chronic conditions targeted by the APM and if initial treatment had demonstrated that their condition could be controlled effectively through a standard treatment and care management regimen.

The patient would choose a Treatment Team that participates in the APM to provide ongoing treatment and care management for the condition. The Treatment Team might or might not be the same as the Diagnostic Team that provided initial treatment for the patient. For example, a physician practice specializing in the chronic condition might have diagnosed and provided initial treatment for the condition (supported by a Diagnosis and Initial Treatment Payment), but the patient might then choose to receive ongoing treatment for the condition from their primary care physician. A patient might have received diagnosis and initial treatment in one community but will receive their ongoing treatment in a different community and will need to find a new Treatment Team there.

Before a patient designated the Treatment Team to provide services, the Team would describe the services that it would deliver and the standards for service delivery that it committed to meet. The Team could also ask the patient to commit to actions that would support good outcomes from treatment. In particular, the Team could ask the patient to only obtain treatment services related to their condition from the members of the Team unless the Team specifically recommends that the patient receive services from other providers.

ii. Payments to the Treatment Team

The Treatment Team would receive a single, pre-defined bundled quarterly Treatment and Care Management Payment to support all of the services required for treatment of the chronic condition and management of the patient's care for that condition. The payment would be expected to cover the costs of office visits and other patient contacts, tests and imaging studies, and any procedures performed by the members of the Team over a three-month period.

Physician practices on the Treatment Team would not bill or be paid for office visits or other traditional Evaluation & Management services. Revenues would come only through the quarterly bundled payment. If the patient receives treatment services from providers other than the Treatment Team during the three-month period in which a Treatment and Care Management Payment is in effect, all or part of the payments the payer makes to those providers would be deducted from the Treatment and Care Management Payment.

Similar to the Diagnosis and Initial Treatment Payment, the costs of any medications prescribed for treatment would not be included in the bundled payment; they would be paid for separately by the patient or the pa-

tient's insurance plan. Laboratory tests or imaging studies that are very expensive and only needed in certain circumstances would also be paid for separately. If the medications are purchased and administered by the Treatment Team, or if the tests/studies are performed by the Treatment Team, the separate payment would be designed to cover the out-of-pocket costs incurred by the Team (e.g., the acquisition cost of the medication or the materials required for tests).

Patients would be stratified into three categories – Low Need/Risk, Moderate Need/Risk, and High Need/Risk – based on characteristics that affect the time or costs of delivering evidence-based treatment or care management or that affect the ability to achieve desirable outcomes. Payments would be higher for patients in categories that require more time or more services. For example, payments might be higher for patients with more severe symptoms or other health problems that require additional time or services.

A Treatment Team would be responsible for dividing the Treatment and Care Management Payment among the Team members to cover the costs they incur in delivering specific types of services to the patient. For example, a Treatment Team might consist of a primary care practice and a specialty physician practice; the primary care practice would provide most of the direct services to the patient, but it would consult with the specialty practice as needed to ensure that the most appropriate treatments are being used and to revise treatment plans when the patient's circumstances change. The primary care practice could bill the patient's health insurance plan for the Treatment and Care Management Payment each quarter, and then use a portion of that Payment to pay the specialty practice a quarterly retainer fee for the patient.

iii. Accountability for Utilization and Spending

The Treatment Team would be held accountable for utilization and spending in three ways:

- **Bundled Payment:** The structure of the Treatment and Care Management Payment would make the Treatment Team directly accountable for utilization and spending on all planned services related to treatment other than the out-of-pocket costs of medications and of infrequent, expensive tests.
- **Outcome-Based Payment:** The Treatment Team would be accountable for avoiding exacerbations of the chronic condition that require an emergency department visit or hospitalization. If a patient in the Low Need/Risk or Moderate Need/Risk categories visits the ED or is hospitalized during a calendar quarter, the Treatment Team would not receive a Treatment and Care Management Payment for that patient in that quarter. If a patient in the High Need/Risk category visits the ED or is hospitalized, the Treatment and Care Management Payment would be reduced by a pre-defined percentage (e.g., 25%). The amounts of the Treatment and Care Management Payment for each category of patients would be set based on the costs of delivering services and the expected rates of ED visits/hospitalizations in each category.

- **Evidence-Based Care:** The Treatment Team would be required to follow evidence-based clinical standards or guidelines in determining which tests, medications, and procedures to perform or order. If the Team failed to follow the guidelines for a patient and did not document the reason for deviating from the guidelines, it would not receive the Treatment and Care Management Payment for that patient in that three-month period.

iv. Accountability for Quality and Outcomes

The Treatment Team would be held accountable for quality and outcomes in two ways:

- **Evidence-Based Care Standards:** In addition to defining which tests, medications, and procedures were appropriate, the evidence-based clinical standards or guidelines would also define any other services or methods of delivery of services that had been demonstrated to result in better treatment outcomes for patients. If the Team failed to follow the guidelines for a patient and did not document the reason for deviating from the guidelines, or if the Team failed to meet the service standards that it had committed to meet when the patient chose it to deliver care, the Team would not receive the Treatment and Care Management Payment for that patient.
- **Desirable Patient-Reported Outcomes:** One or more patient-reported outcome measures would be defined that are relevant to the specific chronic condition being treated. The Treatment and Care Management Payment would be reduced by a pre-defined amount for an individual patient when a desirable outcome did not occur for that patient or when an undesirable outcome did occur. In addition, for chronic conditions where effective treatment can slow the progression of the condition, the Treatment Team could receive a bonus payment for each patient that did not progress to a higher level of severity.

v. Patient Cost-Sharing

The patient would be responsible for paying a fixed quarterly copayment for the services supported by the Treatment and Care Management Payment that are delivered by the Treatment Team or by providers approved by the Treatment Team. This copayment would be set at a level that is at or below the total of the cost-sharing amounts that the patient might expect to pay currently for individual services they would receive as part of treatment for their chronic condition.

If the patient receives treatment services from other providers without approval from the Treatment Team during the period of time that the Treatment and Care Management Payment is in effect, the patient would pay additional cost-sharing for those services.

c. Continued Treatment for Patients with a Difficult-to-Control Condition

i. Eligibility of Patients and Designation of the Treatment Team

Patients would be eligible to receive services supported by the APM in this phase of care if they have been diagnosed with one of the chronic conditions targeted by the APM and if standard treatments and care management regimens were not controlling the patient's symptoms effectively or if special treatments were needed that required more intensive supervision. Patients might become eligible for this category of care after having received care for their condition for a period of time if the condition worsened or if the patient developed other health problems that made the condition more difficult to manage. Patients might also "graduate" from this category and move to the "Well-Controlled Condition" category if a new type of treatment was developed that worked more effectively or if another health problem was resolved.

The patient would choose a Treatment Team that participates in the APM to provide ongoing treatment and care management for the condition. The Treatment Team might or might not be the same as the Diagnostic Team that provided initial treatment or a Treatment Team that previously provided treatment for the patient. For example, a primary care practice might only provide treatment and care management for patients in the well-controlled category, and refer a patient to a different physician practice that specializes in the chronic condition if the patient's condition becomes more difficult to control. A specialty physician practice might treat both types of patients or decide to focus solely or primarily on the patients with more difficult-to-control conditions.

Before a patient designated the Treatment Team to provide services, the Team would describe the services that it would deliver and the standards for service delivery that it committed to meet. The Team could also ask the patient to commit to actions that would support good outcomes from treatment. In particular, the Team could ask the patient to only obtain treatment services related to their condition from the members of the Team unless the Team specifically recommends that the patient receive services from other providers.

ii. Payments to the Treatment Team

The Treatment Team would receive a single, pre-defined bundled quarterly Treatment and Care Management Payment to support all of the services required for treatment of the chronic condition and management of the patient's care for that condition. The payment would be expected to cover the costs of office visits and other patient contacts, tests and imaging studies, and procedures performed by the members of the Team to treat the patient during a three-month period of time.

Similar to the Diagnosis and Initial Treatment Payment and the Treatment and Care Management Payment for patients with a well-controlled condition, physician practices on the Treatment Team for a patient with a difficult-to-control patient would not bill or be paid for office

visits or other traditional Evaluation & Management services; revenues would only come through the quarterly bundled Treatment and Care Management Payment. If the patient receives treatment services from providers other than the Treatment Team during the three months in which a Treatment and Care Management Payment is in effect, all or part of the payments the payer makes to those providers would be deducted from the Treatment and Care Management Payment.

Similar to the other payments, the costs of any medications prescribed for treatment would not be included in the bundled payment, but would be paid for separately by the patient or the patient's insurance plan. Laboratory tests, imaging studies, or procedures that are very expensive and only needed in certain circumstances would also be paid for separately. If the medications are purchased and administered by the Treatment Team, or if the tests/studies/procedures are performed by the Treatment Team, the separate payment would cover the out-of-pocket costs incurred by the Team (e.g., the acquisition cost of the medication or of the materials required for tests).

Patients would be stratified into three categories – Moderate Need/Risk, High Need/Risk, and Very High Need/Risk – based on characteristics that affect the time or costs of delivering evidence-based treatment or care management or that affect the Team's ability to achieve desirable outcomes for the patient. Payments would be higher for patients in categories that require more time or more services. For example, payments would be higher for patients with more severe symptoms or other health problems that require additional time or services.

iii. Accountability for Utilization and Spending

The Treatment Team would be held accountable for utilization and spending in three ways:

- **Bundled Payment:** The structure of the Treatment and Care Management Payment would make the Treatment Team directly accountable for utilization and spending on all planned services related to treatment other than the out-of-pocket costs of medications and infrequent, expensive tests.
- **Outcome-Based Payment:** The Treatment Team would be accountable for avoiding exacerbations of the chronic condition that require an emergency department visit or hospitalization. If a patient in the Moderate Need/Risk or High Need/Risk categories visits the ED or is hospitalized during a calendar quarter, the Treatment Team would not receive a Treatment and Care Management Payment for that patient in that quarter. If a patient in the Very High Need/Risk category visits the ED or is hospitalized, the Treatment and Care Management Payment would be reduced by a pre-defined percentage (e.g., 25%). The amounts of the Treatment and Care Management Payment for each category of patients would be set based on the costs of delivering services and the expected rates of ED visits/hospitalizations in each category.
- **Evidence-Based Care:** The Treatment Team would be required to follow evidence-based clinical guidelines in determining which tests, medications, and procedures to perform or order. If the Team failed to follow

the guidelines for a patient and did not document the reason for deviating from the guidelines, it would not receive the Treatment and Care Management Payment for that patient.

iv. Accountability for Quality and Outcomes

The Treatment Team would be held accountable for quality and outcomes in two ways:

- **Evidence-Based Care Standards:** In addition to defining which medications and tests were appropriate, the evidence-based clinical standards or guidelines would also define any other services or methods of delivery of services that had been demonstrated to result in better treatment outcomes for patients. If the Team failed to follow the guidelines for a patient and did not document the reason for deviating from the guidelines, or if the Team failed to meet the service standards that it had committed to meet when the patient chose it to deliver care, the Team would not receive the Treatment and Care Management Payment for that patient.
- **Desirable Patient-Reported Outcomes:** One or more patient-reported outcome measures would be defined that are relevant to the specific chronic condition being treated. The Treatment and Care Management Payment would be reduced by a pre-defined amount for an individual patient when a desirable outcome did not occur for that patient or when an undesirable outcome did occur. In addition, for chronic conditions where effective treatment can slow the progression of the condition, the Treatment Team could receive a bonus when a patient did not progress to a higher level of severity.

v. Patient Cost-Sharing

The patient would be responsible for paying a fixed quarterly copayment for the services supported by the Treatment and Care Management Payment that are delivered by the Treatment Team or by providers approved by the Treatment Team. This copayment would be set at a level that is at or below the total of the cost-sharing amounts that the patient might expect to pay currently for individual services they would receive as part of treatment for their chronic condition.

If the patient receives treatment services from other providers without approval from the Treatment Team during the period of time that the Treatment and Care Management Payment is in effect, the patient would pay additional cost-sharing for those services.

d. Hospitalization for an Exacerbation of the Chronic Condition

i. Eligibility Criteria for Patients

Patients would be eligible to receive services supported by the APM in this phase of care if they have been diagnosed with the chronic condition and make a visit to a hospital Emergency Department or are admitted to the hospital for symptoms related to their chronic condition or problems that are determined to be due primarily to that chronic condition or the treatments being used.

ii. Payments to the Hospital

Hospitals would receive three separate types of payment to cover the costs of their services to patients with the chronic condition:

- Standby capacity payments;
- Bundled/warranted payments for ED visits and hospital admissions; and
- Outlier payments.

Standby Capacity Payment

The hospital(s) in the community where the Treatment Teams are located would receive a standard, pre-defined Standby Capacity Payment on a quarterly basis for each patient who is receiving Treatment and Care Management Services from a Treatment Team. A higher amount would be paid for patients in higher need/risk categories. The revenues from these payments would be designed to support the cost of maintaining minimum ED and inpatient capacity at the hospital(s) to address exacerbations of the condition when they occur.

The hospitals would determine the amount of the quarterly Standby Capacity Payment by (1) calculating the minimum fixed cost each hospital would have to incur on a quarterly basis to provide minimum staff and equipment for its ED and inpatient services (i.e., the cost that it would incur if it had only one patient), (2) multiplying that fixed cost by the proportion of the hospital's total patients who come to the hospital for exacerbations of the chronic condition, and (3) dividing the product by the estimated total number of patients in the community with the chronic condition. The amount that any hospital would receive would be smaller if there were more hospitals providing services in the community.

Bundled/Warranted Payment for ED Visits/Hospital Admissions

If a patient with the chronic condition who was receiving services supported by Treatment and Care Management Payments went to a hospital ED or was admitted to the hospital, the hospital would receive a single, standard, pre-defined Chronic Condition Hospital Care Payment to support (1) all of the services the patient needed from the hospital and (2) any post-acute care services needed in the 30 days following the visit or admission that were not being provided by the patient's Treatment Team, such as a stay in a skilled nursing facility, home health services, or a hospital readmission. The hospital would

be responsible for dividing the revenues from the Chronic Condition Hospital Care Payments among any providers who were involved in the patient's care during this phase, including the physicians who would manage the patient's care in the hospital, the skilled nursing facility if the patient received services there, etc. The hospital would not charge for or receive any additional payments for any services delivered to patient, unless the circumstances qualified for an Outlier Payment.

A higher Chronic Condition Hospital Care Payment amount would be paid to the hospital for a patient classified in the higher need/risk categories for the Treatment and Care Management Payments. The amount of the Chronic Condition Hospital Care Payment for a patient in a particular need/risk category would be based on the average additional cost the hospitals would incur for a patient beyond the fixed costs supported by the Standby Capacity Payments, except for the services or costs that would be covered by outlier payments. The expected cost would be determined by estimating (1) the cost for patients who visit the ED but are not admitted to the hospital, (2) the cost for patients who are admitted but do not need post-acute care, and (3) the cost for patients who are admitted and need post-acute care or require a hospital readmission after discharge, and weighting those estimated costs by the estimated percentage of patients that could be expected to need those different combinations of services.

The Chronic Condition Hospital Care Payment would be significantly lower than the typical amount a hospital would receive for an inpatient admission because (1) the payments would be designed for patients who only needed care in the ED as well as those who needed an inpatient admission, and (2) the hospital would also be receiving Standby Capacity Payments.

Outlier Payments

The hospital could receive an Outlier Payment in addition to the Standby Capacity Payment and Chronic Condition Hospital Care Payment for a patient who:

- experienced an unavoidable event during the ED visit or hospital admission that occurs infrequently but typically requires a significant number of additional services or additional time or costs; or
- had unusual characteristics that required additional services or additional time or costs in the delivery of typical services during the ED visit or hospital admission.

For events that occur infrequently but require predictable responses, the hospital would receive a standard, pre-defined Outlier Payment. For unusual events, there would not be a pre-defined payment; instead, the amount of the Outlier Payment would be based on the additional costs that the hospital incurred in delivering care to the patient. The hospital would calculate the actual costs it incurred for the patient's care, and subtract the payments it had otherwise received; the Outlier Payment would be equal to 90% of that amount.

iii. Accountability for Utilization and Spending

The hospital would be held accountable for utilization and spending in two ways:

- **Bundled Payment:** The structure of the Standby Capacity Payment and Chronic Condition Hospital Care Payment would make the hospital directly accountable for utilization and spending on all planned services related to hospital and post-acute care.
- **Warranted Payment:** The hospital would be accountable for avoiding any complications resulting from the hospital treatment that require an emergency department visit or hospitalization, since there would be no additional payment for any additional ED visits, hospital readmissions, etc. during the 30 days following an admission.

iv. Accountability for Quality and Outcomes

The hospital would be held accountable for quality and outcomes in two ways:

- **Evidence-Based Care Standards:** The hospital would be expected to follow evidence-based clinical standards or guidelines that had been demonstrated to result in better treatment outcomes for patients. If the hospital failed to follow the guidelines for a patient and did not document the reason for deviating from the guidelines, it would not receive the Chronic Condition Hospital Care Payment for that patient.
- **Mortality:** In order to ensure that the hospital is not undertreating patients, the rate of death among the patients will be measured during the 30-day period following the ED visit or hospital admission. If the rate of mortality in one or more patient categories increased by a statistically significant amount, the hospital's Standby Capacity and Chronic Condition Hospital Care Payments would be reduced.

v. Patient Cost-Sharing

The patient would be responsible for paying a fixed copayment when a hospital bills for a Chronic Condition Hospital Care Payment. This copayment would be set at a level that is at or below the average total of the cost-sharing amounts that the patient might expect to pay currently for an ED visit or hospital admission.

e. Palliative Care for an Advanced Condition

i. Eligibility of Patients and Designation of Palliative Care Team

Patients would be eligible to receive services supported by the APM in this phase of care if they have been diagnosed with the chronic condition and if the condition has progressed to the point where the patient is experiencing significant pain, rapid functional decline, or other symptoms that would benefit from palliative care services in addition to treatment for the chronic condition itself.

The patient would choose a Palliative Care Team that participates in the APM to provide palliative care for the

condition in addition to or instead of treatment. The Palliative Care Team might or might not be the same as the Treatment Team. For example, a large multi-specialty physician practice might serve as both the Treatment Team and Palliative Care Team, providing both types of services, whereas a small primary care practice or a physician practice specializing in treatment of the condition might serve as the Treatment Team and the patient would choose a separate Palliative Care Team, such as a hospice and palliative care services agency, to provide palliative care.

Before a patient designated the Palliative Care Team to provide services, the Team would describe the services that it would deliver and the standards for service delivery that it committed to meet. The Team could also ask the patient to commit to actions that would support the ability of the Team to most effectively address the patient's palliative care needs. In particular, the Team could ask the patient to only obtain palliative care services related to their condition from the members of the Team unless the Team specifically recommends that the patient receive services from other providers.

ii. Payments to the Palliative Care Team

The Palliative Care Team would receive a single, pre-defined bundled monthly Palliative Care Payment to support all of the services required for palliative care. The payment would be expected to cover the costs of home visits and other patient contacts and services performed by the members of the Team during the month. Services ordinarily expected to be provided for treatment of the chronic condition and management of the patient's care for that condition would not be included unless the patient is no longer receiving treatment for the condition. If the patient receives palliative care services from providers other than the Palliative Care Team during the month in which the Palliative Care Payment is in effect, all or part of the payments the payer makes to those providers would be deducted from the Palliative Care Payment.

Similar to the other payments, the costs of any medications prescribed for palliative care would not be included in the bundled payment, but would be paid for separately by the patient or the patient's insurance plan. If the medications are purchased and administered by the Palliative Care Team, the separate payment would be designed to cover the out-of-pocket costs incurred by the Team to acquire the medication.

Patients would be stratified into four categories – Low Need, Moderate Need, High Need, and Hospice – based on the severity of the patient's symptoms and other characteristics that affect the time or costs of delivering evidence-based palliative care services. Payments would be higher for patients in categories that require more time or more services.

Payments for patients in the Hospice category could be based on the payments currently made for patients eligible for hospice care. In particular, the Palliative Care Team would be expected to pay directly for any ED visits or hospitalizations for patients who are in the Hospice category.

iii. Accountability for Utilization and Spending

The Palliative Care Team would be held accountable for utilization and spending in three ways:

- **Bundled Payment:** The structure of the Palliative Care Payment would make the Palliative Care Team directly accountable for utilization and spending on all planned palliative care services other than the out-of-pocket costs of medications.
- **Hospital Care:** For patients who are not in the Hospice category, the patient's Treatment Team (not the Palliative Care Team) would be accountable for avoiding exacerbations of the chronic condition that require an emergency department visit or hospitalization, not the Palliative Care Team. (The Treatment Team would have the option of contracting with a Palliative Care Team or serving as the Palliative Care Team itself in order to share accountability for avoiding hospitalizations with the palliative care providers.) For patients in the Hospice category, the costs of hospital services would be included in the bundled payment for the Palliative Care Team.
- **Evidence-Based Care:** The Palliative Care Team would be required to follow evidence-based clinical guidelines in determining which palliative care medications and services to deliver or order. If the Team failed to follow the guidelines for a patient and did not document the reason for deviating from the guidelines, the Team would not receive the Palliative Care Payment for that patient during the month.

iv. Accountability for Quality and Outcomes

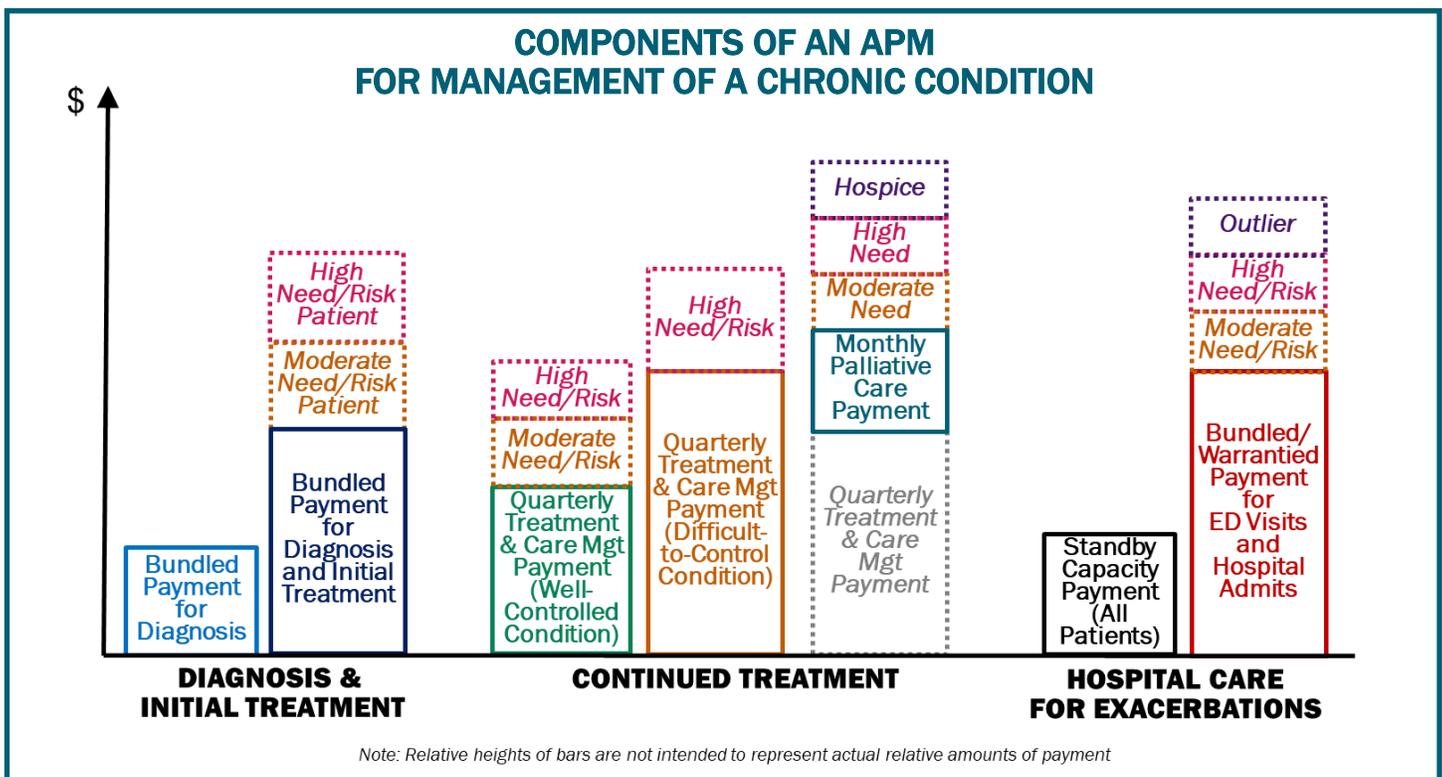
The Palliative Care Team would be held accountable for quality and outcomes in two ways:

- **Evidence-Based Care Standards:** In addition to defining which medications and tests were appropriate, the evidence-based clinical standards or guidelines would also define any other services or methods of delivery of services that had been demonstrated to result in better outcomes for patients. If the Palliative Care Team failed to follow the guidelines for a patient and did not document the reason for deviating from the guidelines, or if the Team failed to meet the service standards that it had committed to meet when the patient chose the Team to deliver palliative care, the Team would not receive the Palliative Care Payment for that patient for that month.
- **Desirable Patient-Reported Outcomes:** One or more patient-reported outcome measures would be defined that are relevant to the chronic condition being treated. The Palliative Care Payment would be reduced by a pre-defined amount for an individual patient when a desirable outcome did not occur for that patient or when an undesirable outcome did occur.

v. Patient Cost-Sharing

The patient would be responsible for paying a fixed co-payment each month for the services supported by the Palliative Care Payment that are delivered by the Palliative Care Team or by providers approved by the Team.

If the patient receives palliative care services from other providers without approval from the Palliative Care Team during the period of time that the Palliative Care Payment is in effect, the patient would pay additional cost-sharing for those services.



5. Operationalizing the APM

In order for Diagnosis Teams, Treatment Teams, Palliative Care Teams and hospitals to be paid for services under the APM, they would submit claims forms for each eligible patient using a series of new codes. Penalties for failure to achieve Desirable Patient-Reported Outcomes would be operationalized as “withholds,” i.e., the default amount of payment for a service code would be calculated by including the maximum penalty for failure to achieve the outcomes, and then additional codes would be created to enable the Team to recoup the penalty when one or more Desirable Outcomes were actually achieved.

Diagnosis and Initial Treatment Payments

- CC011: evaluation of a patient who has not been previously diagnosed with the chronic condition, and who has symptoms of the chronic condition, but is determined not to have the chronic condition
- CC012: initial treatment of a patient newly diagnosed with the chronic condition who is in the low-need/complexity category
- CC013: initial treatment of a patient newly diagnosed with the chronic condition who is in the moderate-need/complexity category
- CC014: initial treatment of a patient newly diagnosed with the chronic condition who is in the high-need/complexity category
- CC015-CC018: additional payments for achieving Desirable Patient-Reported Outcomes
- CC019: maximum additional payment for achieving Desirable Outcomes. If the Diagnosis Team had achieved multiple Desirable Outcomes, it would submit individual codes (CC015-CC018) for each of those outcomes, and if the total additional payments for those codes exceeded the maximum additional payment per patient, the Team would also submit code CC019 and the payment would be made for that code instead of the others. (All of the codes would still be submitted so it was clear which outcomes had been achieved.)

Treatment and Care Management Payments – Well-Controlled Conditions

- CC021: three months of treatment for a patient with a well-controlled condition who meets the criteria for the Low Need/Risk category
- CC022: three months of treatment for a patient with a well-controlled condition who meets the criteria for the Moderate Need/Risk category
- CC023: three months of treatment for a patient with a well-controlled condition who meets the criteria for the High Need/Risk category
- CC025-CC028: additional payments for achieving Desirable Patient-Reported Outcomes
- CC029: maximum additional payment for achieving Desirable Outcomes. If the Treatment Team had achieved multiple Desirable Outcomes, it would sub-

mit individual codes (CC024-CC028) for each of those outcomes, and if the total additional payments for those codes exceeded the maximum additional payment per patient, the Team would also submit code CC029 and the payment would be made for that code instead of the others. (All of the codes would still be submitted so it was clear which outcomes had been achieved.)

Treatment and Care Management Payments – Difficult-to-Control Conditions

- CC031: three months of treatment for a patient with a difficult-to-control condition who meets the criteria for the Moderate Need/Risk category
- CC032: three months of treatment for a patient with a difficult-to-control condition who meets the criteria for the High Need/Risk category
- CC033: three months of treatment for a patient with a difficult-to-control condition who meets the criteria for the Very High Need/Risk category
- CC034-CC038: additional payments for achieving Desirable Patient-Reported Outcomes
- CC039: maximum additional payment for achieving Desirable Outcomes.

Chronic Condition Hospital Care Payments and Outlier Payments

- CC041: hospital care for a patient receiving Treatment and Care Management services in the Well-Controlled Phase and the Low Need/Risk category
- CC042: hospital care for a patient receiving Treatment and Care Management services in the Well-Controlled Phase and the Moderate Need/Risk category
- CC043: hospital care for a patient receiving Treatment and Care Management services in the Well-Controlled Phase and the High Need/Risk category
- CC044: hospital care for a patient receiving Treatment and Care Management services in the Difficult-to-Control Phase and the Moderate Need/Risk category
- CC045: hospital care for a patient receiving Treatment and Care Management services in the Difficult-to-Control Phase and the High Need/Risk category
- CC046: hospital care for a patient receiving Treatment and Care Management services in the Difficult-to-Control Phase and the Very High Need/Risk category
- CC047-CC049: outlier payments

Palliative Care Payments

- CC051: one month of palliative care services for a patient in the Low Need category
- CC052: one month of palliative care services for a patient in the Moderate Need category

- CC053: one month of palliative care services for a patient in the High Need category
- CC054: one month of palliative care services for a patient receiving hospice services
- CC055-CC058: additional payments for achieving Desirable Patient-Reported Outcomes
- CC059: maximum additional payment for achieving Desirable Outcomes.

Submission of Claims

The date of service on the claim would be the last day of the month or quarter in which the services were delivered.

Submission of a claim form for a patient with one of these billing codes would represent a certification by the Team or hospital that:

- The patient met the eligibility criteria for the APM and for the assigned Need/Risk category.
- The patient had received services that met all required evidence-based standards or guidelines for that phase and month or quarter of care.
- The patient had not visited an ED or been admitted to the hospital during the quarter covered by the payment (for those payments that are contingent on avoiding use of hospital services).

If a Team wished to charge patients more than the amount that would be paid by their health plans, the Team would publish its charge for each of the billing codes, and the patient would agree to those charges at the time that the patient was enrolling to receive services from the Team. A single Team would charge the same amount to all of the Team's patients, regardless of their health insurance plan, and the Team would bill the patient for the difference between the charge and the amount paid by the patient's insurance plan.

On a quarterly basis, each Team would calculate its performance on all of the quality measures (both Evidence-Based Care measures and Desirable Patient-Reported Outcome measures). These measures would be calculated separately for patients in each of the need/risk categories. The measure data would be provided to the team's patients and to the health insurance plans for those patients.

The Team would make information on its performance on the quality measures and its charges for services publicly available so that patients seeking a Team could compare the cost and performance of different Teams.

Identifying Chronic Disease-Related ED Visits and Hospital Admissions

A Treatment Team would only be eligible to receive a quarterly Treatment and Care Management Payment for a patient if the patient did not visit an ED and was not admitted to a hospital during the quarter for an exacerbation of the chronic disease. Periodically (e.g., either monthly or quarterly), the patient's health insurance plan would determine whether it had received a claim from a

hospital for a Chronic Condition Hospital Care Payment and a claim from a Treatment Team for a Treatment and Care Management Payment for the same patient during the same quarter, and if so, it would reject payment or request a refund for the Treatment and Care Management Payment.

In order to ensure that patients who made ED visits or had hospital admissions related to the chronic condition were being billed for properly, a hospital participating in the APM would submit to a periodic audit of medical records and claims forms by an independent entity to determine whether patients were being correctly coded.

Payment and Withholds for Reconciliation

If a Diagnosis Team, Treatment Team, or Palliative Care Team submitted a billing code on a claim form, the payer would immediately pay the Team 90% of the pre-defined payment amount assigned to that billing code. The remaining 10% would be held back for a period of 60 days to determine if any claims from other providers were submitted for similar services to the same patients; if so, the total amount withheld would be reduced by the payments made to those providers, and the balance would then be paid to the Team.

Hospital Standby Capacity Payments

Because the hospitals participating in the APM would receive a Standby Capacity Payment for a patient receiving services supported by the APM regardless of whether the patient was actually admitted to the hospital or visited the ED, it would be difficult for the hospital to bill directly for all of these payments. Instead, since the payments would be made if and only if a patient was receiving services supported by Treatment and Care Management Payments, the submission of a claim by a Treatment Team to a participating health insurance plan for a Treatment and Care Management payment would also automatically trigger a Standby Capacity Payment from the health insurance plan to each participating hospital.

The amount of the Standby Capacity Payment should be higher for patients classified in categories that have a higher risk of exacerbations that can lead to ED visits and hospital admissions, so different Standby Capacity Payments should be associated with each of the different codes listed above for different types of patients.

To distinguish the payment made to the Treatment Team from the Standby Capacity Payment made to a hospital, a modifier would be added to the codes listed earlier:

- -OP: Treatment and Care Management Payment to a Treatment Team
- -IP: Standby Capacity Payment to a hospital

For example, if a Treatment Team submits a claim with a CC022 code for a well-controlled, medium need/risk patient, the health plan would issue a payment to the Treatment Team with the amount assigned to the CC022-OP code and modifier, and the health insurance plan would also issue a payment to each participating

hospital with the amount assigned to the CC022-IP code and modifier.

Since the Standby Capacity Payments to hospitals would be tied to claims submitted by Treatment Teams, this means that if the patient is admitted to the hospital and the Treatment Team does not submit a claim for a Treatment and Care Management Payment, the hospital would not receive a Standby Capacity Payment for that patient during that quarter. In this situation, the hospital could submit its own claim for that patient with the modifier -IP attached, since the hospital would know that the patient was participating in the APM and that the hospital had not received a Standby Capacity Payment for that patient.

6. Implementing the APM

a. Obtaining Participation by Payers, Providers, and Patients

The APM would have a number of advantages for payers, providers, and patients that should encourage payers to implement the APM, encourage providers to participate in the APM, and encourage eligible patients to seek care from providers who are participating in the APM.

i. Advantages for Payers

- Participating health insurance plans could reduce spending on plan members who have one or more of multiple types of chronic conditions.
- Participating health insurance plans could eliminate prior authorization programs for medications and procedures, since participating providers would be accountable for following evidence-based treatment guidelines.
- Health insurance plans could implement the APM by creating new billing codes in their existing claims payment system.

ii. Advantages for Providers

- Participating physician practices would have the flexibility to deliver services to their patients in the ways that are most feasible for the practice and most effective for their patients, including office visits, phone calls, and emails with a physician or clinician, and visits and calls with nurses and other types of staff.
- Participating physician practices would receive higher payments to cover the additional time they would spend with patients with greater needs.
- Participating physician practices would be held accountable for whether a patient they had explicitly enrolled for services had visited an ED or was hospitalized for an exacerbation of the chronic condition the practice had committed to manage. The practice would not be held accountable for the total cost of the hospitalization or for other services the patient is receiving from the practice or from other providers. The practice would know in advance what rate of hospitalizations it would be expected to achieve for its patients.

- Participating physician practices would be responsible for following evidence-based clinical guidelines, but would not be penalized for delivering care that their patients needed nor would they be penalized for increases in the amounts that other providers charged for their services or for increases in the prices of drugs and medical devices.
- Participating physician practices would know when to expect payment and how much to expect based on the bills they submit to payers and the cost-sharing charged to patients. The largest financial loss the practice could experience would be the loss of the payments under the APM.
- Physician practices could charge more for their service if they could deliver better outcomes that patients were willing to pay more for.
- Participating hospitals would no longer have all of their revenues tied to the number of patients admitted to the hospital; the hospital could support efforts to reduce hospital admissions and readmissions without losing money by doing so.
- Hospice agencies and other palliative care providers could deliver palliative care services to patients who needed them without requiring the patient to give up treatment services.
- Participating physician practices and hospitals could bill for services using their standard billing systems.

iii. Advantages for Patients

- Patients would have the choice of whether to receive the services supported by the APM based on a clear understanding of what services they would receive, the actions they would need to take, and the results they could expect to achieve.
- Patients could choose different teams of providers in different phases of their care needs, and they could change to different teams multiple times if they wished to do so.
- Patients would know that their physician would be rewarded for helping the patient avoid exacerbations of their chronic condition but would have no financial incentive to withhold needed care.
- Patients would know how much they would need to pay for the services before choosing to receive them.
- Participating patients would experience fewer severe symptoms from their chronic disease. They would receive more care at home and require fewer visits to emergency departments and fewer admissions to hospitals to treat severe symptoms.
- Patients would have the ability to compare the performance and prices of different Diagnostic Teams, Treatment Teams, and Palliative Care Teams in order to choose the Teams they would use.

b. Finalizing the APM Parameters

A “beta test” of the APM will likely be needed with willing providers in order to finalize several key parameters of the APM:

- **Criteria defining the categories of need/risk.** The categories should be defined so that they distinguish which patients will be at higher risk of exacerbations and which patients will need more time and care management services in order for a Team to follow evidence-based care guidelines, to avoid hospitalizations, and to improve patient outcomes. However, data may not be available on all of the factors that would be expected to affect need and risk, and the APM will need to be implemented first in order to enable those data to be collected.
- **Dollar amounts of the various payments.** The payment amounts in each phase of care and for each need/risk level should be based on the cost of the services that would be delivered to patients in that phase and level, but the cost of the services will depend on the number of patients a participating Team could manage and the number of patients in each of the need/risk categories, and this can only be estimated after the services are actually implemented with support from the APM.

- **Benchmark rates of condition-related ED visits and hospital admissions.** The performance targets and payment amounts will depend on the benchmark (baseline) rates of ED visits and hospital admissions in each need/risk category, but this can only be determined after actual patients are classified into the need/risk categories.
- **Benchmark rates of desirable outcomes.** Data are not currently being collected for many types of desirable outcomes for chronic diseases because there is no means of paying for the costs of doing so. Consequently, performance targets and payment amounts for many types of desirable outcomes can only be determined after services under the APM begin.

Best estimates of these parameters would be used to initiate the beta test process, and the participating Teams would gather and share data from their actual experience in implementing care changes with payments under the APM in order to make adjustments to the parameters.

ENDNOTES

1. This report will focus on the problems with *current* versions of fee-for-service payment systems. As will be discussed in detail in the report, some of the problems with current systems are not *inherent* to a fee-for-service payment system, but are simply problems with the way payers have chosen to implement fee-for-service payment systems. In addition, there are other problems with the way some payers have implemented fee-for-service payments that will not be addressed here. For example, lack of price transparency about individual services (i.e., the inability to find out the actual amounts that providers are paid by health plans to deliver services) and price discrimination (i.e., the fact that different patients and payers pay different amounts for the same service from the same provider) are common problems with the way fee-for-service payment has been implemented by many private health plans, but these problems do not exist in the fee-for-service payment system used by Medicare. In contrast, the four problems described in the text exist in most Medicare payment systems as well as in the systems used by private payers.
2. H.R. 2, Pub.L. 114–10. The term “alternative payment methodology” is applied to certain payment systems used for Federally Qualified Health Centers and Rural Health Clinics, but this is not an “alternative payment model” as defined in MACRA.
3. Regulations issued by the Centers for Medicare and Medicaid Services (CMS) require that, beginning in 2019, if clinicians are receiving the payments under the APM, 75% of the clinicians must be using Certified Electronic Health Record Technology (CEHRT) to “document and communicate clinical care to their patients or other health care providers,” and if a hospital is receiving the payments, the hospital must be using CEHRT for that purpose. 42 CFR §414.1415(a).
4. Regulations issued by CMS define “comparable” to mean that at least one of the quality measures on which payment in the APM is based “must have an evidence-based focus, be reliable and valid, and meet at least one of the following criteria: (i) be used in the quality performance category under the Merit-Based Incentive Payment System (MIPS); (ii) be endorsed by a consensus-based entity, (iii) developed by CMS as part of its Measure Development Plan; (iv) submitted in response to the MIPS Call for Quality Measures; or (v) determined by CMS to have an evidence-based focus and to be reliable and valid.” In addition, the quality measures used by the APM must include at least one outcome measure unless CMS determines there are no available or applicable outcome measures. 42 CFR §414.1415(b).
5. 42 U.S.C. 1315a.
6. The Health Care Innovation Awards were grants made by CMMI to 146 projects in 2012 and 2014. Congress presumably excluded them from the definition of an APM because they were innovative service delivery projects supported by one-time grants, not payment models designed for continuation and replication. More information about the Health Care Innovation Awards is available at <https://innovation.cms.gov/initiatives/index.html#views=Health%20Care%20Innovation%20Awards>
7. 42 U.S.C. 1395jjj. The full set of provisions of Section 1899 is available at https://www.ssa.gov/OP_Home/ssact/title18/1899.htm.
8. To date, CMS has only used this authority to implement separate “tracks” that require ACOs to take “downside risk” (i.e., requiring them to pay CMS if Medicare spending increases beyond expected amounts), but the law permits a broader range of payment models to be used. It is not clear that the “downside risk” tracks that have been implemented meet the requirements of Section 1899, since the provision for “other payment models” requires that spending cannot be higher for an individual ACO, but CMS only requires ACOs to repay a portion of any increase in spending.
9. 42 U.S.C. 1395cc-3.
10. Information on the demonstrations implemented under Section 1866C are available at <https://innovation.cms.gov/initiatives/Medicare-Health-Care-Quality/>.
11. CMS has implemented the Bundled Payments for Care Improvement (BPCI) and Bundled Payments for Care Improvement – Advanced APMs, which it believes satisfy the Congressional requirement to establish a National Pilot Program on Payment Bundling.
12. The statutory requirements for the Independence at Home Demonstration are at 42 U.S.C. 1395cc-5. Information on how CMS has implemented the requirements is available at: <https://innovation.cms.gov/initiatives/Independence-at-Home/>.
13. As of 2018, many of the APMs that had been implemented by CMS did not meet the requirements of either the statute or the CMS regulations to qualify as an Advanced APM. For example, the vast majority of ACOs participating in the Medicare Shared Savings Program are in “Track 1,” which is an APM but does not qualify as an Advanced APM.
14. 42 U.S.C. 1395ee (c).
15. 42 CFR §414.1465.
16. More information on the proposals submitted to the Physician-Focused Payment Model Technical Advisory Committee and the recommendations PTAC has made to the Secretary of Health and Human Services is available at <https://aspe.hhs.gov/proposal-submissions-physician-focused-payment-model-technical-advisory-committee>.

17. 42 CFR §414.1410 -1445.
18. The regulation defines this limit in terms of the average revenues of all entities participating in the APM, not the revenues of each individual entity.
19. As of 2018, no medical home programs had been expanded under Section 1115A.
20. *Alternative Payment Models in the Quality Payment Program as of February 2018*. Centers for Medicare and Medicaid Services (2018). Available at: <https://gpp-cm-prod-content.s3.amazonaws.com/uploads/113/Comprehensive-List-of-APMs.pdf>
21. *CMS Multi-Payer Other Payer Advanced APMs in the Quality Payment Program for Performance Year 2019*. Centers for Medicare and Medicaid Services (2018). Available at <https://www.cms.gov/Medicare/Quality-Payment-Program/Resource-Library/Medicaid-Other-Payer-Advanced-APM-determination-list.pdf>
22. Information about the Merit-Based Incentive Payment System is available at: <https://gpp.cms.gov/mips/overview>.
23. 42 CFR §414.1370.
24. Miller HD. *How to Fix the Medicare Shared Savings Program*. Center for Healthcare Quality and Payment Reform (June 2018). Available at: http://www.chqpr.org/downloads/How_to_Fix_the_Medicare_Shared_Savings_Program.pdf
25. Lewin Group. *CMS Bundled Payments for Care Improvement Initiative Models 2-4: Year 5 Evaluation & Monitoring Annual Report*. CMS (October 2018). Available at: <https://downloads.cms.gov/files/cmimi/bpci-models2-4-yr5evalrpt.pdf>
26. *Was the Medicare Shared Savings Program Successful in 2017?* Center for Healthcare Quality and Payment Reform (August 2018). Available at: <http://chqpr.org/blog/index.php/2018/08/was-the-medicare-shared-savings-program-successful-in-2017/>
27. Miller HD. *Why Value-Based Payment Isn't Working, and How to Fix It*. Center for Healthcare Quality and Payment Reform (October 2017). Available at: <http://www.chqpr.org/downloads/WhyVBPIsNotWorking.pdf>
28. *Ibid.*
29. 42 U.S.C. 1315a.
30. For simplicity, in this report, the term “provider” will be used to refer to any individual or organization that delivers healthcare services, including physicians, nurse practitioners, physician assistants, hospitals, skilled nursing facilities, home health agencies, hospice programs, etc.
31. The term “total cost of care” is typically used to describe the total amount of payments made to healthcare providers for all of the services delivered to a patient or to a group of patients during a period of time. It is the “cost” to those who pay for the services, not the cost that the providers actually incur in delivering the services. As will be discussed later in the report, the cost of a service for payers (i.e., the amount they pay for it) is often very different than the cost providers incur in delivering the service.
32. Miller HD. *A Better Way to Pay for Cancer Care*. Center for Healthcare Quality and Payment Reform. Available at: <http://www.chqpr.org/downloads/BetterPaymentforCancerCare.pdf>.
33. Miller HD. *Why Value-Based Payment Isn't Working, and How to Fix It*. Center for Healthcare Quality and Payment Reform (October 2017). Available at: <http://www.chqpr.org/downloads/WhyVBPIsNotWorking.pdf>.
34. 42 U.S.C. 1315a.
35. Round 2 of the Health Care Innovation Awards program created by the Center for Medicare and Medicaid Innovation (CMMI) provided grants for innovative projects and also explicitly required that “All applicants must submit, as part of their application, the design of a payment model that is consistent with the new service delivery model.” <https://innovation.cms.gov/initiatives/Health-Care-Innovation-Awards/faq-round-2.html>
36. It is possible that improvements in the patient health problems that lead to unplanned care could be accompanied by worsening of other things that are undesirable for the patient but do not have a direct or comparable impact on healthcare spending. For example, a drug that prevents post-treatment infections but causes severe pain or mental confusion for a patient might result in lower healthcare spending but lower quality of life for the patient.
37. Fleming-Dutra KE, et al. “Prevalence of Inappropriate Antibiotic Prescriptions Among US Ambulatory Care Visits, 2010-2011,” *JAMA* 315(17):1864-1873 (2016).
38. Shiley KT, et al. “The Use of Antimicrobial Agents After Diagnosis of Viral Respiratory Tract Infections in Hospitalized Adults: Antibiotics or Anxiolytics?” *Infect Control Hosp Epidemiol* 31(11): 1177-1183 (November 2010).
39. Rogg J, et al. “The Frequency and Cost of Redundant Laboratory Testing for Transferred Emergency Department Patients,” *Annals of Emergency Medicine* 60(48): S13 (October 2012).
40. Chen CL, Clay TH, McLeod S. “A Revised Estimate of Costs Associated with Routine Preoperative Testing in Medicare Cataract Patients with a Procedure-Specific Indicator,” *JAMA Ophthalmol.* 136(3):231-238 (2018).
41. Flynn TW, Smith B, Chou R. “Appropriate Use of Diagnostic Imaging in Low Back Pain: A Reminder That Unnecessary Imaging May Do as Much Harm as Good.” *Journal of Orthopaedic & Sports Physical Therapy* 41 (11): 838-846 (November 2011).
42. <http://www.ChoosingWisely.org>
43. <https://www.uspreventiveservicestaskforce.org/>
44. Brown D, Clement F. “Calculating Health Care Waste in Washington State,” *JAMA Internal Medicine* (August 6, 2018).
45. Yang CL, et al. “Misdiagnosis of Asthma in Schoolchildren,” *Pediatric Pulmonology* 52:293-302 (2017)
46. Aaron SD, et al. “Reevaluation of Diagnosis in Adults with Physician-Diagnosed Asthma,” *JAMA* 317(3): 269-279 (2017).
47. Heffler, E et al. “Misdiagnosis of Asthma and COPD and Underuse of Spirometry in Primary Care Unselected Patients,” *Respiratory Medicine* 142:48-52 (2018).
48. In fact, the default treatment for COPD can cause death if used for asthma patients.

49. Society for Cardiovascular Angiography and Interventions. *Five Things Physicians and Patients Should Question*. Available at: <http://www.choosingwisely.org/wp-content/uploads/2015/02/SCAI-Choosing-Wisely-List.pdf>
50. Patel MR, et al. "Low Diagnostic Yield of Elective Coronary Angiography," *N Engl J Med* 362:886-95 (2010).
51. Fox J, Haig AJ, Todey B, Challa S. "The Effect of Required Physiatrist Consultation on Surgery Rates for Back Pain," *SPINE* 38(3): E178-E184 (February 2013).
52. Benatar S, Garrett AB, Howell E, Palmer A. "Midwifery Care at a Freestanding Birth Center: A Safe and Effective Alternative to Maternity Care," *Health Services Research* 48(5):1750-1768 (October 2013).
53. Montalto M, Lui B, Mullins A, Woodmason K. "Medically-Managed Hospital in the Home: 7 Year Study of Mortality and Unplanned Interruption," *Australian Health Review* 34:269-275 (2010).
54. Higgins A, Veselovskiy G, Schinkel J. "National Estimates of Price Variation by Site of Care," *Am J Manag Care* 22 (3): e116-e121 (2016).
55. Blot K, et al. "Prevention of Central Line-Associated Bloodstream Infections Through Quality Improvement Interventions: A Systematic Review and Meta-analysis," *Clinical Infectious Diseases* 59(1):96-105 (2014).
56. HHS Office of Inspector General, *Adverse Events in Rehabilitation Hospitals: National Incidence Among Medicare Beneficiaries* (July 2016). HHS Office of Inspector General, *Hospital Incident Reporting Systems Do Not Capture Most Patient Harm* (January 2012). HHS Office of Inspector General, *Adverse Events in Hospitals: National Incidence Among Medicare Beneficiaries* (November 2010).
57. Psaltopoulou T, Ilias I, Alevizaki M. "The Role of Diet and Lifestyle in Primary, Secondary, and Tertiary Diabetes Prevention: A Review of Meta-Analyses." *The Review of Diabetic Studies* 7(1): 26-35 (2010).
58. Spitalnic P. *Certification of the Medicare Diabetes Prevention Program*. CMS Office of the Actuary (March 14, 2016). Available at: <https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/ActuarialStudies/Downloads/Diabetes-Prevention-Certification-2016-03-14.pdf>.
59. Siu AL. "Screening for High Blood Pressure in Adults: U.S. Preventive Services Task Force Recommendation Statement." *Annals of Internal Medicine* 163(10):778-787 (November 2015).
60. Wilkinson TMA, et al. "Early Therapy Improves Outcomes of Exacerbations of Chronic Obstructive Pulmonary Disease," *Am J Respir Crit Care Med* 169:1298-1303 (2004).
61. NORC. "Innovative Oncology Business Solutions, Inc." in *HCIA Disease-Specific Evaluation – Third Annual Report* (February 2016). Available at: <https://downloads.cms.gov/files/cmmi/hcia-diseasespecific-thirdannualrpt.pdf>
62. Johnson DS, Kapanian T, Taylor R, Meyer KB. "Going Upstream: Coordination to Improve CKD Care," *Seminars in Dialysis* 29(2): 125-134 (2016).
63. *Medicare Claims Processing Manual*. Centers for Medicare and Medicaid Services.
64. Committee on Diagnostic Error in Health Care. *Improving Diagnosis in Health Care*. The National Academies Press (2015).
65. Petty TL. "Benefits of and Barriers to the Widespread Use of Spirometry," *Curr Opin Pulm Med* 11:115-120 (2005).
66. Shannon RP, et al. "Using Real-Time Problem Solving to Eliminate Central Line Infections," *Joint Commission Journal on Quality and Patient Safety* 32(9): 479-487 (2006).
67. As noted earlier, it could be desirable to combine this particular change in care delivery with another change such that the combined savings exceeds the combined cost. However, this may not always be possible.
68. Bai G, Anderson GF. "US Hospitals Are Still Using Chargemaster Markups to Maximize Revenues," *Health Affairs* 35(9): 1658-1664 (2016).
69. If the amount that the physician practice pays for a drug depends on the volume of drugs that it purchases (e.g., if the supplier of the drug offers a quantity discount or rebate of some kind), then the drug will still represent a variable cost to the practice, but the variable cost amount may also vary based on the number of times the practice treats patients using the drug.
70. Komisar HL, Feder J. *Transforming care for Medicare beneficiaries with chronic conditions and long-term care needs: coordinating care across all services*. Scan Foundation and Georgetown University. October 2011. Available at: http://www.thescanfoundation.org/sites/default/files/Georgetown_Transforming_Care.pdf.
71. Miller HD. *Making the Business Case for Payment and Delivery Reform*. Robert Wood Johnson Foundation. Available at <http://www.chqpr.org/downloads/BusinessCaseforPaymentReform.pdf>
72. As noted in Section III.C, if the service delivery approach improves outcomes but requires higher spending, payers could make a decision to pay providers in a way that encourages or requires that approach, but this would not be an "Alternative Payment Model" as defined in MACRA.
73. O'Connor AM, Llewellyn-Thomas HA, Barry Flood A. "Modifying Unwarranted Variations in Health Care: Shared Decision Making Using Patient Decision Aids." *Health Affairs* 63(1) (2004)
74. Oliverez M. "Chronic Care Management Reimbursement: Why Aren't More Doctors Billing For It?" *MD Magazine* (January 31, 2017).
75. 42 U.S.C. 1395w-4.
76. *Medicare Claims Processing Manual*. Centers for Medicare and Medicaid Services.

77. The fact that average costs increase and profits decrease with lower volume and vice versa is not unique to healthcare; the same rule holds in most industries. In manufacturing, when more products are sold, the fixed cost of the manufacturing plant and equipment can be spread across a larger number of products, reducing the average cost of each product and increasing the manufacturer's profits. As the cost decreases, the manufacturer can choose to reduce the price, encouraging even more sales and even higher profits. In manufacturing, this is a virtuous cycle. In healthcare, it is not, because success in healthcare is keeping people healthy so they need fewer healthcare services. In manufacturing, a fixed price for each product rewards the manufacturer with higher profits for successfully selling more products, but in healthcare, the fixed fee for a service penalizes a healthcare provider for keeping a patient healthy.
78. Schurrer J, et al. *Evaluation of the Diffusion and Impact of the Chronic Care Management (CCM) Services: Final Report*. Mathematica Policy Research (November 2, 2017).
79. Ward JC, et al. "Impact on Oncology Practices of Including Drug Costs in Bundled Payments," *Journal of Oncology Practice* 14(5):e259-e268 (May 2018).
80. Institute of Medicine Committee to Study the Prevention of Low Birthweight. *Preventing Low Birthweight*. National Academy Press (1985).
81. *Medicare Claims Processing Manual*. Centers for Medicare and Medicaid Services.
82. Miller HD. *Measuring and Assigning Accountability for Healthcare Spending*. Center for Healthcare Quality and Payment Reform (2014). Available at: <http://www.chqpr.org/downloads/AccountabilityforHealthcareSpending.pdf>
83. In the 2019 HCC Risk Adjustment Model, a patient's risk score is increased by .302 if they have colorectal cancer and a patient's risk score is increased by .305 if they have diabetes with acute or chronic complications. Centers for Medicare and Medicaid Services. *Advance Notice of Methodological Changes for Calendar Year (CY) 2019 for the Medicare Advantage (MA) CMS-HCC Risk Adjustment Model* (December 27, 2017).
84. In regulations for the Inpatient Prospective Payment System, CMS stated, "Because the DRGs were developed to group clinically similar patients, an extremely important means of communication between the clinical and financial aspects of care was created. DRGs provided administrators and physicians with a meaningful basis for evaluating both the process of providing care and the associated financial impacts. Development of care pathways by DRG and profit-and-loss reports by DRG product lines became commonplace. With the adoption of these new management methods, length of stay and the use of ancillary services dropped dramatically...The vast majority of modifications to the DRGs since the inception of the Medicare inpatient hospital prospective payment system ... have almost always been the result of clinicians identifying specific types of patients with unique needs...Central to the success of the Medicare inpatient hospital prospective payment system is that DRGs have remained a clinical description of why the patient required hospitalization." 66 Federal Register 22668, May 4, 2001.
85. *Medicare Claims Processing Manual*. Centers for Medicare and Medicaid Services.
86. Centers for Medicare and Medicaid Services. *BPCI-Advanced Clinical Episode Construction Specifications: Model Years 1 and 2*. (March 2018).
87. As discussed under Option 9, Medicare will also pay an additional "outlier" payment for some patients who need to stay in the hospital for an unusually long period of time or who need unusually intense services during their stay.
88. A standby capacity payment that is based on the population of the community being served or a group of patients with a particular health condition or risk factor would be very different than what is commonly being described as a "population-based payment." The standby capacity payment would only be designed to cover the fixed cost of a service that needs to be on standby, not the full cost of that service nor the cost of services that do not have to be standing by at all times, whereas "population-based payments" are generally intended to cover the costs of services that are actually delivered.
89. *Medicare Claims Processing Manual*. Centers for Medicare and Medicaid Services.
90. Another example of a standby payment in healthcare is the payment for vaccines made through the Center for Disease Control's Strategic National Stockpile program. See <https://www.cdc.gov/cpr/stockpile/index.htm> .
91. 42 U.S.C. 1395ww and 42 CFR §412.101.
92. *Medicare Claims Processing Manual*. Centers for Medicare and Medicaid Services.
93. Kalman NS, et al. "Removing a Constraint on Hospital Utilization: A Natural Experiment in Maryland," *Am J Manag Care* 20(6): e191-e199 (2014).
94. de Lagasnerie G, et al. "Tapering Payments in Hospitals: Experiences in OECD Countries," *OECD Health Working Papers* No. 78 (2015)
95. *Medicare Claims Processing Manual*. Centers for Medicare and Medicaid Services.
96. Adamski J, et al. "Risk Sharing Arrangements for Pharmaceuticals: Potential Considerations and Recommendations for European Payers," *BMC Health Services Research* 10:153 (2010).
97. For example, in the Total Patient Revenue system in Maryland, a hospital bills for each service that is delivered to each patient using the state-approved charge in effect at the time. If the total revenue from all of the services delivered exceeds the hospital's budget during the course of the year, it is required to reduce its charges to patients during the subsequent year in order to offset the excess revenues generated the previous year.
98. *Medicare Claims Processing Manual*. Centers for Medicare and Medicaid Services.
99. *Medicare Claims Processing Manual*. Centers for Medicare and Medicaid Services.
100. *Medicare Claims Processing Manual*. Centers for Medicare and Medicaid Services.
101. United States Government Accountability Office. *Payment Methods for Certain Cancer Hospitals Should Be Revised to Promote Efficiency* (February 2015).

102. Linesch SJ. "The Coded DOFR and Its Role in Accountable Care," *CAPG Health* (Summer 2015).
103. Centers for Medicare and Medicaid Services. "Behavioral Health Integration Services." *MLN Fact Sheet* ICN 909432 (January 2018).
104. Centers for Medicare and Medicaid Services. *Participation in the Acute Care Episode (ACE) Demonstration Minimum Requirements – Formation of a Physician-Hospital Organization (PHO)*. (June 2008). Available at <https://innovation.cms.gov/Files/x/Physician-Hospital-Organization-Requirements.pdf>
105. Miller HD. *Bundling Better: How Medicare Should Pay for Comprehensive Care*. Center for Healthcare Quality and Payment Reform. <http://www.chqpr.org/downloads/BundlingBetter.pdf>
106. Lewin Group. *CMS Bundled Payments for Care Improvement Initiative Models 2-4: Year 5 Evaluation & Monitoring Annual Report*. CMS (October 2018). Available at: <https://downloads.cms.gov/files/cmimi/bpci-models2-4-yr5evalrpt.pdf>.
107. Hoff T. "The Battle of the Bundle: Lessons from My Mother's Partial Hip Replacement," *Health Affairs* 36(8):1511-1514. (2017)
108. Centers for Medicare and Medicaid Services. *Updated Guidance for Long-Term Care (LTC) Facility Participation in the Initiative to Reduce Avoidable Hospitalizations Among Nursing Facility Residents – Payment Reform*. (February 22, 2017).
109. Lucia K, Hoadley J, Williams A. *Balance Billing by Health Care Providers: Assessing Consumer Protections Across States*. The Commonwealth Fund (June 2017).
110. Zhang H, Cowling DW, Facer M. "Comparing the Effects of Reference Pricing and Centers-of-Excellence Approaches to Value-Based Benefit Design." *Health Affairs* 36 (12): 2094-2101 (2017).
111. For example, if the APM pays for a new type of post-acute home rehabilitation service for knee replacement patients, and the goal of that service is to reduce the utilization of inpatient rehabilitation services following knee surgery, then the APM needs a definition of the specific type of inpatient rehabilitation services that are expected to be reduced. The APM for knee surgery would not be expected to reduce inpatient rehabilitation services for head injuries or substance use disorders, so it would be inappropriate to hold the providers in the APM accountable for all inpatient rehabilitation services the patients receive.
112. It is possible that an increase in the number of knee replacement procedures could result in improved outcomes for the patients, but unless it also generated savings in healthcare costs, it would not qualify as an APM.
113. RTI International and Actuarial Research Corporation. *OCM Performance-Based Payment Methodology, Version 3.2* (December 27, 2017).
114. Condition-based composite spending measures can be defined for both acute conditions and chronic diseases. Although spending measures for chronic disease care are also sometimes referred to as "episode measures," the fact that the patient generally needs care on an ongoing basis until death means there really is no true "episode" that has beginning and end points, so the "episodes" used in APMs for chronic disease management are merely arbitrary periods of time, such as a month or a year.
115. Although the episodes may have similar names in the different episode grouper systems, there are important differences in how the episodes are defined and how the groupers determine which services to include. Comparisons of different episode groupers have shown that the differences in the methodologies can result in large differences in the way services and spending are assigned to episodes. For example, a 2006 study by the Medicare Payment Advisory Commission found that two commonly used episode groupers, when applied to the same population of Medicare patients, calculated significantly different amounts of spending in episodes with similar names. (Medicare Payment Advisory Commission. *June 2006 Report to the Congress*.) A 2008 study conducted by Acumen, LLC for the Centers for Medicare and Medicaid Services found that one of these episode groupers assigned the majority of a sample patient's spending to a Pneumonia episode, whereas the other grouper assigned the majority of the patient's spending to an Alzheimer's Disease episode. (MaCurdy T et al. *Evaluating the Functionality of the Symmetry ETG and Medstat MEG Software in Forming Episodes of Care using Medicare Data*. Acumen LLC August 2008). A 2012 study conducted for the U.S. Bureau of Economic Analysis found that those same two episode groupers, when applied to a group of commercially insured patients, produced very different classifications of spending into episodes. (Rosen A et al. "Comparing Commercial Systems for Characterizing Episodes of Care." *U.S. Bureau of Economic Analysis Working Paper Series* June 2012.)
116. Delia D. "Spending Carveouts Substantially Improve the Accuracy of Performance Measurement in Shared Savings Arrangements: Findings from Simulation Analysis of Medicaid ACOs." *INQUIRY: The Journal of Health Care Organization, Provision, and Financing* 54(17): 1-11 (2017). The study examined the effects of excluding three types of uncontrollable spending – spending on injury care, spending on custodial care in facilities, and spending above \$100,000 per patient – from a measure of total spending and found that excluding these types of spending dramatically reduced bias and error in estimates of savings due to the actions of the ACO.
117. de Brantes F, D'Andrea G, Rosenthal MB. Should healthcare come with a warranty? *Health Aff* 28(4): w678-w687 (2009).

118. National Quality Forum Endorsed Measures #0704 (Proportion of patients hospitalized with AMI that have a potentially avoidable complication), #0705 (Proportion of patients hospitalized with stroke that have a potentially avoidable complication), #0708 (Proportion of patients hospitalized with pneumonia that have a potentially avoidable complication), and #0709 (Proportion of patients with a chronic condition that have a potentially avoidable complication during a calendar year). Available at <http://www.qualityforum.org>.
119. Goldfield N, Kelly WP, Patel K. "Potentially preventable events: an actionable set of measures for linking quality improvement and cost savings." *Q Manage Health Care* 21(4):213-219. (October-December 2012).
120. Centers for Medicare and Medicaid Services. *BPCI-Advanced Clinical Episode Reconciliation Specifications: Model Years 1 and 2* (June 2018).
121. Centers for Medicare and Medicaid Services. *CMS Standardization Methodology for Allowed Amount – v.7*. (June 2018).
122. HealthPartners. *Total Care Relative Resource Value (TCRRV) White Paper* (September 2017).
123. Patient-specific appropriateness measures can be used when the evidence is clear as to what is and is not appropriate for specific patients, whereas population-based measures should be used if the evidence is primarily based on population-level correlations. Lawson EH, et al. "Appropriateness Criteria to Assess Variations in Surgical Procedure Use in the United States." *Archives of Surgery* 146(12):1433-1440 (2011). Shekelle PG, et al. "The Reproducibility of a Method to Identify the Overuse and Underuse of Medical Procedures." *New England Journal of Medicine* 338(26):1888-1895 (1998).
124. RTI International and Actuarial Research Corporation. *OCM Performance-Based Payment Methodology, Version 3.2* (December 27, 2017).
125. Graham KL, et al. "Differences Between Early and Late Readmissions Among Patients: A Cohort Study," *Ann Intern Med* 162:741-749 (2015).
126. Centers for Medicare and Medicaid Services. *BPCI-Advanced Clinical Episode Reconciliation Specifications: Model Years 1 and 2* (June 2018).
127. Centers for Medicare and Medicaid Services. *BPCI-Advanced Clinical Episode Reconciliation Specifications: Model Years 1 and 2* (June 2018).
128. RTI International and Actuarial Research Corporation. *OCM Performance-Based Payment Methodology, Version 3.2* (December 27, 2017).
129. Kiefe, CI et al. "Improving Quality Improvement Using Achievable Benchmarks for Physician Feedback: A Randomized Controlled Trial," *JAMA* 285:2871-2879 (2001)
130. Centers for Medicare and Medicaid Services. *BPCI-Advanced Clinical Episode Reconciliation Specifications: Model Years 1 and 2*. (June 2018)
131. Centers for Medicare and Medicaid Services. *Medicare Program; Medicare Shared Savings Program: Accountable Care Organizations Final Rule*. 76 FR 67802 (November 2, 2011).
132. RTI International and Actuarial Research Corporation. *Independence at Home Demonstration Actuarial Shared Savings Methodology Specifications* (September 2017).
133. This is because complete claims data will generally not be available for several months after the end of the year, it will then take time to calculate the Benchmark based on those data, and the APM participant will want to know the Target for the coming year well before the year begins.
134. Problems with data lags can be mitigated somewhat if Targets and Performance Measures can be defined for shorter periods of time, e.g., 3 months rather a full year, since that enables more recent data to be used. However, Targets and measures defined over short periods of time can be less reliable for smaller providers since they will be based on fewer patients or services, less reliable for all providers if there is significant seasonal variation in patient needs, and less accurate if savings are expected to occur over a longer period of time than the timeframe measured.
135. Centers for Medicare and Medicaid Services. *Medicare Shared Savings Program: Shared Savings and Losses and Assignment Methodology Specifications* (May 2018).
136. Miller HD. *How to Fix the Medicare Shared Savings Program*. Center for Healthcare Quality and Payment Reform (June 2018). Available at: http://www.chqpr.org/downloads/How_to_Fix_the_Medicare_Shared_Savings_Program.pdf
137. Section 1899(d)(1)(B) of the Social Security Act (42 U.S.C. 1395jjj).
138. Section 1115A(b)(3)(B) of the Social Security Act states "The Secretary shall terminate or modify the design and implementation of a model unless the Secretary determines...that the model is expected to – (i) improve the quality of care...without increasing spending...; (ii) reducing spending...without reducing the quality of care; or (iii) improve the quality of care and reduce spending." (42 U.S.C. 1315a).
139. Adjusting a spending measure based on relative levels of patient risk is equivalent to estimating what the spending level would have been for a "risk-neutral" population, i.e., where every patient has the same level of risk. Instead of defining the Benchmark for a risk-neutral population, some APMs adjust the Benchmark for an APM participant based on the actual risk scores of the patients who received care from the APM participant. Under this approach, the actual utilization or spending by the APM participant is not risk-adjusted at all, and the adjustments to the Benchmark allow it to be treated as the APM participant's "expected" utilization or spending for patients with similar characteristics. However, this makes it more difficult to define a Target performance level prospectively, since it depends on knowing the characteristics of the patients who actually received care from the APM participant during the performance period.

140. In the Medicare Shared Savings Program, when a patient is first assigned to an ACO, CMS calculates that patient's risk score using the standard Hierarchical Condition Category risk adjustment methodology and adjusts the benchmark during the first performance year based on that risk score, i.e., if the patient has a high risk score, the benchmark will be adjusted upward to reflect the expectation that spending will be higher for that patient. However, for patients who were first assigned to the ACO in previous years, the benchmark is adjusted differently. Even if new diagnoses have been assigned to the patients, the benchmark will not be increased on that basis, but if the existing patients' health status improves, the benchmark will be reduced. Centers for Medicare and Medicaid Services. *Medicare Shared Savings Program: Shared Savings and Losses and Assignment Methodology Specifications* (May 2018).
141. Dobson DaVanzo & Associates LLC. *Estimates of Savings by Medicare Shared Savings Program Accountable Care Organizations* (August 2018).
142. CMS is using a different methodology for trending when it recalculates an ACO's benchmark in second and subsequent agreement periods.
143. 42 CFR Part 510.
144. Selective participation could occur either because of choices the patient makes (e.g., if they do not want to receive the services supported by the APM) or because of choices the provider makes (e.g., if the provider discourages higher-risk patients from participating).
145. Information on unusual differences in patients may be easier to obtain in the future since successes of "precision medicine" in treating very specific types of conditions such as cancer have led to more systematic efforts to do genetic testing and other types of testing to identify when rare conditions are present.
146. Centers for Medicare and Medicaid Services. *Medicare Program; Medicare Shared Savings Program: Accountable Care Organizations Final Rule*. 76 FR 67802 (November 2, 2011).
147. Although many APMs truncate spending for unusually high-cost patients, a patient may have an unusually high cost for a particular treatment without exceeding the outlier thresholds typically used.
148. An average of 8 patients (100/12) patients would be treated each month, and 7 would cost \$1,000, but 1 would cost \$10,000, for an average of \$2,125.
149. Yu WW, Machlin S. *Examination of Skewed Health Expenditure Data from the Medical Expenditure Panel Survey (MEPS)*. Agency for Healthcare Research and Quality Working Paper No. 04002 (October 2004).
150. As discussed earlier, a study by Derek DeLia found that the inclusion of uncontrollable components of spending in a shared savings calculation created significant biases and errors in estimates of savings compared to the true savings that the ACO produced. DeLia D. "Spending Carveouts Substantially Improve the Accuracy of Performance Measurement in Shared Savings Arrangements: Findings from Simulation Analysis of Medicaid ACOs." *INQUIRY: The Journal of Health Care Organization, Provision, and Financing* 54(17): 1-11 (2017).
151. DeLia D, Hoover H, Cantor JC. "Statistical Uncertainty in the Medicare Shared Savings Program." *Medicare & Medicaid Research Review* 2(4) (2012).
152. DeLia D. "Leaving It to Chance: The Effects of Random Variation in Shared Savings Arrangements." *Health Services Research Methodology* 13:219-240 (2013).
153. Ash AS, et al. *Statistical Issues in Assessing Hospital Performance*. Committee of Presidents of Statistical Societies (January 27, 2012).
154. Decisions will also need to be made about how to operationalize "paying a penalty to the payer." For example, there might be no transfer of cash from the provider to the payer, but merely a reduction in future payments from the payer to the provider equivalent to the amount of the penalty.
155. In CPC+, the new payments that the practice receives are divided into two components: a "Care Management Fee" and a "Performance-Based Incentive Payment," and the penalty for high rates of ED visits or hospital admissions is proportional to one-half of the Performance-Based Incentive Payment amount. Centers for Medicare and Medicaid Services. *CPC+ Payment Methodologies: Beneficiary Attribution, Care Management Fee, Performance-Based Incentive Payment, and Payment Under the Medicare Physician Fee Schedule* (February 17, 2017).
156. Centers for Medicare and Medicaid Services. *BPCI-Advanced Clinical Episode Reconciliation Specifications: Model Years 1 and 2* (June 2018).
157. RTI International and Actuarial Research Corporation. *OCM Performance-Based Payment Methodology, Version 3.2* (December 27, 2017). As of 2018, there were no practices participating in Track 2 of OCM.
158. Centers for Medicare and Medicaid Services. *BPCI-Advanced Clinical Episode Reconciliation Specifications: Model Years 1 and 2* (June 2018).
159. 42 CFR §414.1415 defines the limits as either "8 percent of the estimated average total Medicare Parts A and B revenues of participating APM Entities" or "3 percent of the expected expenditures for which an APM Entity is responsible under the APM."
160. *Accelerating the Implementation of Value-Based Care and Payment: Recommendations from the 2016 National Payment Reform Summit*. Network for Regional Healthcare Improvement (2016). Available at: http://www.nrhi.org/uploads/2016_nationalpaymentreformsummit.pdf
161. In the Oncology Care Model, CMS stated that the practice would be dropped from the program if it does not achieve savings within 3 years.
162. If the spending on the second measure was simply increasing or decreasing randomly, but the provider's performance on the first measure was consistently poor, then the provider would receive a penalty randomly from year to year, rather than consistently.
163. If the APM is being implemented by a specific payer, that payer could provide the stop-loss protection directly, rather than the APM participant attempting to purchase it separately.

164. Casale AS, et al. "ProvenCareSM: A Provider-Driven Pay-for-Performance Program for Acute Episodic Cardiac Surgical Care," *Annals of Surgery* 246(4): 613-623 (October 2007).
165. Johnson LL, Becker RL. "An Alternative Health-Care Reimbursement System—Application of Arthroscopy and Financial Warranty: Results of a Two-Year Pilot Study." *Arthroscopy* 10(4):462–70 (1994).
166. 42 U.S.C. 1395cc-5.
167. Kline RM, et al. "Centers for Medicare and Medicaid Services: Using an Episode-Based Payment Model to Improve Oncology Care," *Journal of Oncology Practice* 11(2):114-117 (2015).
168. 42 U.S.C. 1315a.
169. Mathematica Policy Research. *Evaluation of the Comprehensive Primary Care Initiative: Fourth Annual Report* (May 2018).
170. As discussed in Section I, MACRA requires that an APM base payment on quality measures "comparable" to those used in the Merit-Based Incentive Payment System. Regulations issued by the Centers for Medicare and Medicaid Services define "comparable" to mean that at least one of the quality measures on which payment in the APM is based "must have an evidence-based focus, be reliable and valid, and meet at least one of the following criteria: (i) be used in the quality performance category under the Merit-Based Incentive Payment System (MIPS); (ii) be endorsed by a consensus-based entity (e.g., the National Quality Forum), (iii) developed by CMS as part of its Measure Development Plan; (iv) submitted in response to the MIPS Call for Quality Measures; or (v) determined by CMS to have an evidence-based focus and to be reliable and valid." In addition, the quality measures used by the APM must include at least one outcome measure unless CMS determines there are no available or applicable outcome measures. 42 CFR §414.1415(b).
171. As an illustration, the rate of new lung cancer diagnoses is about 300 per 100,000 Medicare beneficiaries, which means that an ACO with 10,000 assigned beneficiaries might expect to see an average of 30 new lung cancer cases each year. The different drug treatment regimens for lung cancer vary in cost from \$2,500 to \$105,000 depending on patient characteristics. (Ward JC, et al. "Impact on Oncology Practices of Including Drug Costs in Bundled Payments," *Journal of Oncology Practice* 14 (5):e259-e268.) If the 15 patients who needed the highest-cost regimens were given the lower cost regimens instead, it would reduce overall spending for the ACO by over 1%.
172. MACRA added Section 1848(q)(2)(C) "Emphasizing Outcome Measures Under the Quality Performance Category" to the Social Security Act. It states: "In applying subparagraph (B)(i) [Quality measures], the Secretary shall, as feasible, emphasize the application of outcome measures."
173. Centers for Medicare and Medicaid Services. *Overview of CJR Quality Measures, Composite Quality Score, and Pay-for-Performance Methodology*. Available at <https://innovation.cms.gov/Files/x/cjr-qualsup.pdf>.
174. MacLean CH, Kerr EA, Qaseem A. "Time Out – Charting a Path for Improving Performance Measurement." *N Engl J Med* 378:1757-1761 (May 10, 2018).
175. In the 2006 Medicare Oncology Demonstration, CMS created 81 new G-codes in order to obtain more detailed data for patients being treated for 13 different types of cancer regarding (1) the stage of cancer, (2) the purpose of oncology visits, and (3) whether the treatment being used adhered to clinical guidelines. Participation was voluntary, and a physician who submitted codes in all three categories was paid \$23 in addition to the standard payment for an Evaluation & Management visit. (The physician received a payment of \$7.67 for the code for each category, but no payment was made unless codes for all three categories were submitted.) Centers for Medicare and Medicaid Services. "2006 Oncology Demonstration Project." *MLN Matters* MM4219 (2006).
176. It is important to recognize that *lack of evidence regarding benefits* of delivering a process or achieving an intermediate outcome level is not the same as *evidence that there are no benefits*. In fact, if an APM is designed to support a different approach to care, there may well not be any evidence about the effectiveness of that care delivery approach because it has not previously been feasible to deliver it due to payment barriers, so there may be no way to set an evidence-based threshold. By measuring both processes and outcomes once the APM is being used, it may be possible to develop such evidence, but this would not be available when the APM is first implemented.
177. For example, patients are not expected to pay for the cost of drugs during clinical trials that are used to establish the effectiveness of drugs, but once the drugs are approved, patients who need the drugs may not be able to afford to purchase them.
178. Jaeschke R, Singer J, Guyatt GH. "Measurement of Health Status. Ascertaining the Minimal Clinically Important Difference." *Control Clin Trials* 10(4):407-15 (1989). Angst F, Aeschlimann A, Angst J. "The Minimal Clinically Important Difference Raised the Significance of Outcome Effects Above the Statistical Level, With Methodological Implications for Future Studies." *Journal of Clinical Epidemiology* 82:128-136 (2017).
179. Shoemaker MJ, et al. "Clinically Meaningful Change Estimates for the Six-Minute Walk Test and Daily Activity in Individuals with Chronic Heart Failure." *Cardiopulmonary Physical Therapy Journal* 24(3):21-29 (2013).
180. Caution is needed with Population-Level Targets defined around rare events, because the measure will be unreliable when used with small patient populations. For example, if a quality problem is only expected to occur at a rate of 2 per thousand for patients with a particular condition, then if a provider participating in the APM is only treating 100 such patients per year, the rate that would be calculated for that provider could only be 0 per thousand or at least 10 per thousand, depending on whether the rare event occurred that year or not. Even if the "true" quality performance of the provider was 2 per thousand, that could only be determined over a longer period of time, and in any given year, the measured rate would seem very low (0 per thousand) or extremely high (10 per thousand). It may still be appropriate to use such a measure to evaluate the quality of care for all patients receiving care under the APM as a whole, but it would be inappropriate to use the population-level measure to evaluate the performance of an individual provider.

181. "Improving Quality Improvement Using Achievable Benchmarks for Physician Feedback: A Randomized Controlled Trial," *op cit*.
182. Centers for Medicare and Medicaid Services. *CPC+ Payment Methodologies: Beneficiary Attribution, Care Management Fee, Performance-Based Incentive Payment, and Payment Under the Medicare Physician Fee Schedule* (February 17, 2017).
183. Centers for Medicare and Medicaid Services. *CPC+ Payment Methodologies: Beneficiary Attribution, Care Management Fee, Performance-Based Incentive Payment, and Payment Under the Medicare Physician Fee Schedule* (February 17, 2017).
184. Centers for Medicare and Medicaid Services. *Merit-Based Incentive Payment System (MIPS): Scoring 101 Guide for Year 2 (2018)*. Available at: <https://www.cms.gov/Medicare/Quality-Payment-Program/Resource-Library/2018-MIPS-Scoring-Guide.pdf>.
185. Centers for Medicare and Medicaid Services. *Medicare Shared Savings Program: Shared Savings and Losses and Assignment Methodology Specifications* (May 2018).
186. Centers for Medicare and Medicaid Services. *BPCI-Advanced Clinical Episode Reconciliation Specifications: Model Years 1 and 2* (June 2018).
187. Reeves D, et al. "Combining Multiple Indicators of Clinical Quality: An Evaluation of Different Analytic Approaches," *Medical Care* 45(6):489-496 (June 2007).
188. Centers for Medicare and Medicaid Services. *Merit-Based Incentive Payment System (MIPS): Scoring 101 Guide for Year 2 (2018)*. Available at: <https://www.cms.gov/Medicare/Quality-Payment-Program/Resource-Library/2018-MIPS-Scoring-Guide.pdf>.
189. Centers for Medicare and Medicaid Services. *Medicare Diabetes Prevention Program (MDPP): Quick Reference Guide to Payment and Billing*. (2018) Available at: <https://innovation.cms.gov/Files/x/mdpp-billingpayment-refguide.pdf>.
190. Parsons AS, et al. "Medicare Underpayment for Diabetes Prevention Program: Implications for DPP Suppliers," *Am J Manag Care* 24(10):475-478 (2018).
191. There is growing interest in the use of clinical "pathways" to ensure that patients are receiving the best, evidence-based care. Because there will never be sufficient evidence to definitively recommend what should be done for every patient because of the many differences in patients, it is not expected that 100% of the patients will receive care that is specified in the pathway; indeed, 80% adherence has been used in many cases to define what is "good" adherence to the pathway. Lower adherence rates would be seen as a problem, but higher adherence rates may also be seen as a problem because it could suggest that the pathway guidelines are being applied too simplistically or blindly. In order to use pathway compliance in an outcome-based payment structure, "adherence with the pathway" for the purposes of the measure would need to be defined as either (1) following the pathway recommendations when it is applicable to a patient, or (2) not following the pathway recommendation because there is no relevant recommendation or the applicable recommendation has been specifically determined to be inappropriate for the specific patient. This would enable 100% success to be achieved on the measure. A second measure – percentage of patients where the pathway recommendation was deemed inappropriate – could be defined and tracked in order to identify providers who have unusually high numbers of such patients, but payment would not be contingent on this second measure because it could only be assessed at a population level.
192. Kim S, Cohen MA, and Netessine S. "Performance Contracting in After-Sales Service Supply Chains." *Management Science*. 53(12): 1843-1858 (December 2007). Vitasek K. "The Rolls-Royce Of Effective Performance-Based Collaboration" *MaintenanceTechnology.com* (2012). Available at: <http://www.maintenancetechnology.com/2012/06/the-rolls-royce-of-effective-performance-based-collaboration/>.
193. If the APM is being implemented by a specific payer, that payer could provide the stop-loss protection directly, rather than the APM participant attempting to purchase it separately.
194. Centers for Medicare and Medicaid Services. *BPCI-Advanced Clinical Episode Construction Specifications: Model Years 1 and 2* (March 2018).
195. Centers for Medicare and Medicaid Services. *Medicare Shared Savings Program: Shared Savings and Losses and Assignment Methodology Specifications* (May 2018).
196. Centers for Medicare and Medicaid Services. *CPC+ Payment Methodologies: Beneficiary Attribution, Care Management Fee, Performance-Based Incentive Payment, and Payment Under the Medicare Physician Fee Schedule* (February 17, 2017).
197. Miller HD. *Measuring and Assigning Accountability for Healthcare Spending*. Center for Healthcare Quality and Payment Reform (2014). Available at: <http://www.chqpr.org/downloads/AccountabilityforHealthcareSpending.pdf>.
198. Centers for Medicare and Medicaid Services. *BPCI-Advanced Clinical Episode Construction Specifications: Model Years 1 and 2* (March 2018).
199. RTI International and Actuarial Research Corporation. *OCM Performance-Based Payment Methodology, Version 3.2* (December 27, 2017).
200. The Lewin Group. *CMS Bundled Payments for Care Improvement Initiative Models 2-4: Year 5 Evaluation & Monitoring Annual Report* (October 2018).
201. 42 U.S.C. §1395dd.
202. Casale AS, et al. "'ProvenCareSM': A Provider-Driven Pay-for-Performance Program for Acute Episodic Cardiac Surgical Care," *Annals of Surgery* 246(4): 613-623 (October 2007).
203. Centers for Medicare and Medicaid Services. *Chronic Care Management Services*. CMS Medicare Learning Network ICN 909188 (December 2016).

204. In the May 9, 2016 edition of “CPC+ Frequently Asked Questions,” CMS stated “We have seen in the Original CPC Model that shared savings ...has certain limitations in motivating practices to control total cost of care. For example...total cost of care may be challenging for small primary care practices to control and there are no independent incentives for improved quality; andthe amount of any shared savings payments is unknown in advance and the complexity of the regionally aggregated formula and paucity of actionable cost data leaves practices doubtful of achieving any return. The incentive payment methodology in CPC+ will address some of these limitations. The incentive design is stronger because it can be more closely measured at the practice level, will incorporate measures that primary care practices can directly impact, and will be more easily understood by practice leaders.”
205. Centers for Medicare and Medicaid Services. *Updated Guidance for Long-Term Care (LTC) Facility Participation in the Initiative to Reduce Avoidable Hospitalizations Among Nursing Facility Residents – Payment Reform*. (February 22, 2017).
206. Additional examples of APMs are available in: Marks SS and Miller HD. *A Guide to Physician-Focused Alternative Payment Models*. American Medical Association and Center for Healthcare Quality and Payment Reform (2015). Available at <http://www.chqpr.org/downloads/Physician-FocusedAlternativePaymentModels.pdf> . Miller HD. *The Building Blocks of Successful Payment Reform: Designing Payment Systems That Support Higher-Value Health Care*. Network for Regional Healthcare Improvement. (April 2015). Available at: <http://www.chqpr.org/downloads/BuildingBlocksofSuccessfulPaymentReform.pdf>.
207. Parsons AS, et al. “Medicare Underpayment for Diabetes Prevention Program: Implications for DPP Suppliers,” *Am J Manag Care* 24(10):475-478 (2018).
208. Hospitals use ICD-10-PCS codes to describe procedures for purposes of Medicare billing under the Inpatient Prospective Payment System, whereas hospital charges for services are typically based on CPT/HCPCS codes.
209. CPT is a registered trademark of the American Medical Association (AMA). More information about CPT codes is available at: <https://www.ama-assn.org/practice-management/cpt-current-procedural-terminology>.
210. HCPCS is divided into two principal subsystems, referred to as Level I and Level II of the HCPCS. Level I of the HCPCS consists of CPT Codes. Level II of the HCPCS is a standardized coding system that is used primarily to identify products, supplies, and services not included in the CPT codes, such as ambulance services and durable medical equipment, prosthetics, orthotics, and supplies (DMEPOS) when used outside a physician's office. Because Medicare and other insurers cover a variety of services, supplies, and equipment that are not identified by CPT codes, the Level II HCPCS codes were established for submitting claims for these items. The development and use of Level II of the HCPCS began in the 1980s. Level II codes are also referred to as alpha-numeric codes because they consist of a single alphabetical letter followed by 4 numeric digits, while CPT codes are identified using 5 numeric digits. Centers for Medicare and Medicaid Services. “HCPCS – General Information” at <https://www.cms.gov/Medicare/Coding/MedHCPCSGenInfo/index.html>.
211. For example, Medicare does not pay physicians for CPT Code 99441: “telephone evaluation and management service.”
212. Information about the CPT Editorial Panel is available at <https://www.ama-assn.org/practice-management/cpt-editorial-panel>.
213. Information about the HCPCS Workgroup is available at <https://www.cms.gov/Medicare/Coding/MedHCPCSGenInfo/HCPCSPublicMeetings.html>.
214. If a CPT code is created for the service described in the temporary code, then the CPT code will be treated as the “permanent” code and the temporary code will be eliminated. However, there is no limit on how long a “temporary” code can exist.
215. Whereas all codes are 5 characters long, in CPT codes, all 5 characters are numeric, whereas “G-codes” begin with the letter G followed by 4 numbers, “S-codes” begin with the letter S followed by 4 numbers, and “T-codes” begin with the letter T followed by 5 numbers.
216. Private health plans could also use a G-code if it adequately described the service they wished to pay for.
217. The add-on code lists can be downloaded at <https://www.cms.gov/Medicare/Coding/NationalCorrectCodInitEd/Add-On-Code-Edits.html> .
218. Centers for Medicare and Medicaid Services. *Medicare Diabetes Prevention Program (MDPP): Quick Reference Guide to Payment and Billing*. (2018) Available at: <https://innovation.cms.gov/Files/x/mdpp-billingpayment-refguide.pdf>.
219. For example, in the CMS Bundled Payments for Care Improvement-Advanced APM, patients with chronic diseases are included but only if they are hospitalized and then only for a 90-day period following discharge. A CMS-sponsored white paper describing an “episode payment model” for patients with heart disease proposed to pay the same amount for a patient every month regardless of whether they were newly diagnosed and never treated before, well-managed on standard treatments, refractory to standard treatments, etc.; a higher amount would only be paid if the patient received a revascularization procedure. <http://hcp-ian.org/workproducts/cad-whitepaper-final.pdf>

220. In the Inpatient Prospective Payment System, the hospital does not bill for a specific DRG; the Medicare intermediary calculates the DRG based on the ICD-10 diagnosis codes and ICD-10-PCS procedure codes submitted on the claim form and then pays the hospital based on that DRG. However, it would be feasible to have each hospital calculate the DRG itself and submit a claim with the appropriate DRG on it, and some private payers pay based on the DRG code submitted by the hospital on the claim form.
221. Moreover, the same risk scoring system could be used to create different categories for different services if the risk scores affect costs or outcomes in different ways for the different services.
222. Centers for Medicare and Medicaid Services. *CPC+ Payment Methodologies: Beneficiary Attribution, Care Management Fee, Performance-Based Incentive Payment, and Payment Under the Medicare Physician Fee Schedule* (February 17, 2017).
223. The PTP Column One/Column Two Correct Coding edit file can be downloaded at <https://www.cms.gov/Medicare/Coding/NationalCorrectCodInitEd/NCCI-Coding-Edits.html>.
224. There could also be situations in which the APM code could not be billed if claims for other services were submitted.
225. In these cases, the codes would not be used to submit a claim to a payer, but to determine how the payment should be allocated among the providers who are participating.
226. Centers for Medicare and Medicaid Services. *MACRA Patient Relationship Categories and Codes: Frequently Asked Question (FAQ)*. (May 2018). Available at <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/Patient-Relationship-Categories-and-Codes-webinar-FAQ.PDF>.
227. Centers for Medicare and Medicaid Services. *BPCI-Advanced Clinical Episode Reconciliation Specifications: Model Years 1 and 2* (June 2018).
228. A risk-stratified PMPM payment is effectively the same thing as a set of condition-based payments, where the conditions are defined by the risk categories.
229. Centers for Medicare and Medicaid Services. *Updated Guidance for Long-Term Care (LTC) Facility Participation in the Initiative to Reduce Avoidable Hospitalizations Among Nursing Facility Residents – Payment Reform*. (February 22, 2017).
230. Creation and use of parallel sets of codes in both ICD-10 and CPT/HCPCS would not be unprecedented, however, since CMS currently maintains a detailed set of ICD-10-PCS procedure codes that are used for billing for procedures and services delivered during inpatient hospital stays even though there are also CPT/HCPCS codes describing most of the same procedures and services.
231. Determining a payment amount based on both a CPT/HCPCS code and principal diagnosis code would be equivalent to having a 10-character coding system, with the first 5 alphanumeric characters based on the service (i.e., the CPT/HCPCS code) and the second 5 alphanumeric characters based on a diagnosis or condition (i.e., the ICD-10 code).
232. For example, nephrologists receive monthly payments to care for patients with end-stage renal disease (ESRD) who are receiving hemodialysis. There are four separate categories of CPT codes and different payments based on the age of the patient, and the nephrologist determines which code to use based on the age of the patient and the frequency of services provided to the patient during the month.
233. *Medicare Claims Processing Manual*. Centers for Medicare and Medicaid Services.
234. *Medicare Claims Processing Manual*. Centers for Medicare and Medicaid Services.
235. Some Health Information Exchanges (HIEs) only have participation by a subset of providers in a community, so the information they contain for any patient may be incomplete.
236. More information on APCDs is available at <https://www.apcdcouncil.org/>.
237. The Office of Inspector General in the Department of Health and Human Services criticized the Demonstration for paying too much for the symptom assessment and for failing to provide sufficient guidance to physicians to ensure that their coding of symptoms was consistent. DHHS Office of Inspector General. *Cost and Performance of Medicare’s 2005 Chemotherapy Demonstration Project* (August 2006). The project was terminated after one year and replaced by the 2006 Medicare Oncology Demonstration.
238. *Medicare Claims Processing Manual*. Centers for Medicare and Medicaid Services.
239. Calderwood, MS et al. “Centers for Medicare and Medicaid Services Hospital-Acquired Conditions Policy for Central Line-Associated Bloodstream Infection (CLABSI) and Catheter-Associated Urinary Tract Infection (CAUTI) Shows Minimal Impact on Hospital Reimbursement,” *Infection Control & Hospital Epidemiology* 39(8): 897-901 (August 2018).
240. The diagnosis code may be used to determine whether there will be any payment at all, but not the amount that will be paid.
241. Centers for Medicare and Medicaid Services. “2006 Oncology Demonstration Project.” *MLN Matters* MM4219 (2006)
242. Centers for Medicare and Medicaid Services. 2018 Merit-Based Incentive Payment System (MIPS) Quality Performance Category: Claims Data Submission Fact Sheet. Available at: <https://www.cms.gov/Medicare/Quality-Payment-Program/Resource-Library/2018-Claims-data-submission-fact-sheet.pdf>.
243. For example, CMS has created a “web reporting interface” that large physician groups can use to report quality data, but this requires entering quality data for a sample of patients that are selected by CMS. Centers for Medicare and Medicaid Services. *2018 Merit-Based Incentive Payment System (MIPS): CMS Web Interface Fact Sheet*. Available at: <https://www.cms.gov/Medicare/Quality-Payment-Program/Resource-Library/2018-CMS-Web-Interface-fact-sheet.pdf>.
244. Centers for Medicare and Medicaid Services. “2006 Oncology Demonstration Project.” *MLN Matters* MM4219 (2006).

245. Section 1833(g) of the Social Security Act (42 U.S.C. 1395l). CMS announced that it would be discontinuing use of this system in 2019.
246. Centers for Medicare and Medicaid Services. *Medicare Diabetes Prevention Program (MDPP): Quick Reference Guide to Payment and Billing*. (2018) Available at: <https://innovation.cms.gov/Files/x/mdpp-billingpayment-refguide.pdf>.
247. Centers for Medicare and Medicaid Services. *Medicare Diabetes Prevention Program (MDPP): Quick Reference Guide to Payment and Billing*. (2018) Available at: <https://innovation.cms.gov/Files/x/mdpp-billingpayment-refguide.pdf>.
248. In 2018, 61% of workers with employer-sponsored insurance were enrolled in plans that are either partially or completely self-funded. The Kaiser Family Foundation. *Employer Health Benefits: 2018 Annual Survey*.
249. Centers for Medicare and Medicaid Services. "Medical Loss Requirements Under the Patient Protection and Affordable Care Act." *Federal Register* 76(235): 76574 (December 7, 2011).
250. Mathews AW. "Behind Your Rising Health-Care Bills: Secret Hospital Deals That Squelch Competition." *Wall Street Journal* (September 18, 2018).
251. Section 1848(n)(9)(A) of the Social Security Act required CMS to develop a public episode grouper that could be used in the Medicare program as well as by commercial payers. The Episode Grouper for Medicare was developed by Brandeis University under a contract to CMS, but CMS did not allow the work to be completed and it has not used the Episode Grouper for Medicare in CMS programs. In the Medicare Access and CHIP Reauthorization Act of 2015 (MACRA), Congress required that CMS establish care episode groups which account for an estimated one-half of expenditures under Medicare Parts A and B. Although MACRA required this work to be completed by the end of 2017, CMS had only proposed eight episode groups by the statutory deadline.
252. For example, the Health Care Learning and Action Network created by the Centers for Medicare and Medicaid Services asks health insurance companies to report "total dollars paid to providers through fee-for-service-based shared-risk (linked to quality) payments," with no effort to differentiate whether payments for services are different than under the standard fee-for-service payment system, what the performance measures are or what the targets for performance are, how much of the payments are at risk, etc. Health Care Learning and Action Network. *APM Measurement: Progress of Alternative Payment Models: Methodology and Results Report*. The MITRE Corporation (2018).
253. DeVore BL, Wilson B, Parsons JJ. *Employer-Led Innovation for Healthcare Delivery and Payment Reform: Intel Corporation and Presbyterian Healthcare Services*. Intel Corporation.
254. Christianson JB, Feldman R. "Exporting the Buyers Health Care Action Group Purchasing Model: Lessons from Other Communities." *The Milbank Quarterly* 83(1): 149-176. (2005).
255. Baumgarten A. *Analysis of Integrated Delivery Systems and New Provider-Sponsored Health Plans*. Robert Wood Johnson Foundation (June 2017).
256. Friedberg MW, et al. *Effects of Health Care Payment Models on Physician Practice in the United States*. RAND Corporation and American Medical Association (2018). Available at: https://www.rand.org/pubs/research_reports/RR2667.html.
257. Miller HD. *Measuring and Assigning Accountability for Healthcare Spending*. Center for Healthcare Quality and Payment Reform (2014). Available at: <http://www.chqpr.org/downloads/AccountabilityforHealthcareSpending.pdf>
258. 42 U.S.C. 1395. The law is named after its author, Congressman Pete Stark of California.
259. Designated health services are: clinical laboratory services; physical therapy, occupational therapy, and outpatient speech-language pathology services; radiology and certain other imaging services; radiation therapy services and supplies; durable medical equipment and supplies; parenteral and enteral nutrients, equipment, and supplies; prosthetics, orthotics, and prosthetic devices and supplies; home health services; outpatient prescription drugs; and inpatient and outpatient hospital services.
260. *Issue Brief: Physician Self-Referral*. American Medical Association (2015).
261. 42 U.S.C. 1320a-7b.
262. Leaman K. "State Anti-Kickback Statutes: Where the Action Is." *Health Law & Policy Brief* 2(2):22-27 (2008).
263. 42 U.S.C. 1320a-7b.
264. DHHS Office of Inspector General. *Special Advisory Bulletin: Gainsharing Arrangements and CMPs for Hospital Payments to Physicians to Reduce or Limit Services to Beneficiaries* (July 1999). Available at: <http://oig.hhs.gov/fraud/docs/alertsandbulletins/gainsh.htm>.
265. DHHS Office of Inspector General. Advisory Opinion No. 08-16 (October 7, 2008).
266. *Network for Regional Healthcare Improvement*. Jewish Healthcare Foundation and Pittsburgh Regional Health Initiative (2016). Available at: <http://www.nrhi.org/uploads/nrhi-roots-publication-june-2016.pdf>.
267. Burke T, Cartwright-Smith L, Pereira E, Rosenbaum S. *The Antitrust Aspects of Health Information Sharing by Public and Private Health Insurers*. George Washington University Department of Health Policy (July 2009).
268. L&M Policy Research. *Evaluation of CMMI Accountable Care Organization Initiatives: Advance Payment ACO Final Report* (November 25, 2016).
269. Miller HD. *Why Value-Based Payment Isn't Working, and How to Fix It*. Center for Healthcare Quality and Payment Reform (October 2017). Available at: <http://www.chqpr.org/downloads/WhyVBPIsNotWorking.pdf>.
270. Blumenthal D. "Making It Easier for Hospitals to Participate In, and Succeed Under, Bundled Payments." *JAMA Internal Medicine* (October 22, 2018).

271. Several CMS APMs limit the impact of high-cost patients by “Winsorizing” the spending for those patients, which means that the amount of spending for an individual patient that is counted for the purposes of the APM is capped. Winsorization reduces the impact of an unusually high-cost patient on a spending measure, but it does not adequately adjust for it. The Winsorization threshold is generally very high, so an unusually expensive patient still contributes an unusually high amount of spending to the average for a small provider, just not as much as otherwise.
272. Johnston KJ, et al. “Association Between Patient Cognitive and Functional Status and Medicare Total Annual Cost of Care: Implications for Value-Based Payment.” *JAMA Intern Med.* (September 17, 2018).
Roberts ET, et al. “Assessment of the Effect of Adjustment for Patient Characteristics on Hospital Readmission Rates: Implications for Pay for Performance.” *JAMA Intern Med* (September 17, 2018).
273. Riley GF. “Impact of Continued Biased Disenrollment from the Medicare Advantage Program to Fee-for-Service,” *Medicare & Medicaid Research Review* 2(4): E1-E16 (2012).
274. Centers for Medicare and Medicaid Services. *BPCI – Advanced: Frequently Asked Questions (FAQs)*. (March 2018).
275. For example, the Center for Medicare and Medicaid Innovation (CMMI) made grants to 146 projects in 2012 and 2014 through the Health Care Innovative Awards program. More information about the Health Care Innovation Awards is available at <https://innovation.cms.gov/initiatives/index.html#views=Health%20Care%20Innovation%20Awards>
276. This is typically described as a “difference in differences” analysis.
277. Ziliak ST, McCloskey DN. *The Cult of Statistical Significance*. University of Michigan Press (2008).
278. A typical actuarial evaluation would assign a 100% probability to the cost of any increase in payment, but a low probability to potential savings, resulting in a conclusion that the APM will increase spending.
279. 42 U.S.C. 1315a
280. Kim H, Meath THA, Grunditz JI, Quiñones AR, Ibrahim SA, McConnell KJ. “Characteristics of hospitals exiting the newly voluntary comprehensive care for joint replacement program,” *JAMA Internal Medicine* (October 22, 2018)
281. 42 U.S.C. 1315a.
282. 42 U.S.C. 1315a.
283. Miller HD. *Accelerating Payment Reform in Medicare: How CMS Can Implement More Successful Alternative Payment Models More Quickly*. Center for Healthcare Quality and Payment Reform (November 2017). Available at: <http://www.chqpr.org/downloads/AcceleratingPaymentReforminMedicare.pdf> .
284. 42 U.S.C. 1315a.
285. 42 U.S.C. 1395w-4.
286. Cramton P, Ellermeyer S, Katzman B. “Designed to Fail: The Medicare Auction for Durable Medical Equipment,” *Economic Inquiry* 53(1): 469-485 (2015).
287. Lyles A, Weiner JP, Shore AD, Christianson J, Solberg LI, Drury P. “Cost and quality trends in direct contracting arrangements,” *Health Aff (Millwood)* 21(1):89-102 (2002).
288. 42 CFR Part 510.
289. Miller HD. *Why Value-Based Payment Isn’t Working, and How to Fix It*. Center for Healthcare Quality and Payment Reform (October 2017). Available at: <http://www.chqpr.org/downloads/WhyVBPIsNotWorking.pdf>.
290. Hoff T. “The Battle of the Bundle: Lessons from My Mother’s Partial Hip Replacement.” *Health Affairs* 36 (8):1511-1514 (August 2017).
291. For an overview of the many different Medicare payment systems, see the “Payment Basics” documents developed by the Medicare Payment Advisory Commission at <http://medpac.gov/documents/payment-basics>.
292. Truven Health Analytics. *The Cost of Having a Baby in the United States*. Childbirth Connection, Catalyst for Payment Reform, and Center for Healthcare Quality and Payment Reform (January 2013).
293. Hill I, et al. *Strong Start for Mothers and Newborns Evaluation: Year 5 Project Synthesis*. Center for Medicare and Medicaid Innovation (October 2018).



CENTER FOR
HEALTHCARE
QUALITY &
PAYMENT REFORM

320 Fort Duquesne Boulevard
Suite 20-J
Pittsburgh, PA 15222

www.CHQPR.org
