
# Measuring and Assigning Accountability for Healthcare Spending

## Fair and Effective Ways to Analyze the Drivers of Healthcare Costs and Transition to Value-Based Payment

Harold D. Miller

# CONTENTS

## Measuring and Assigning Accountability for Healthcare Spending

# EXECUTIVE SUMMARY

In an effort to address the high cost of health care, the federal government, commercial health plans, and other organizations are defining and using measures of health care spending for the purposes of rewarding or penalizing physicians, hospitals, and other health care providers, defining provider networks, and encouraging patients to use particular providers. For example, in the Medicare program, payments to individual physicians and hospitals will be increased or decreased based on measures of spending on the healthcare services their patients receive.

In general, the spending measures that are being used are designed to assign accountability to a *single* physician, hospital, or other provider for *all* of the spending on *all* of the health care services received by a patient during a particular period of time, regardless of which physicians, hospitals, or other providers actually delivered those services. Statistical rules are used to *retrospectively* attribute responsibility to an individual physician, physician practice, hospital, or other provider for the spending on all of the services that a patient received during either an "episode of care" or a calendar year. The spending amount attributed to each provider is divided by a "risk score" in an effort to adjust for differences in patients' needs for services. If the risk-adjusted spending attributed to a provider is higher than the risk-adjusted spending for other providers, it is presumed that the provider is delivering services inefficiently.

## PROBLEMS WITH CURRENT APPROACHES TO MEASURING AND ASSIGNING ACCOUNTABILITY FOR HEALTHCARE SPENDING

Although most payers use similar methodologies for attribution, episode definition, and risk adjustment, the results differ significantly depending on the detailed specifications of the methodologies. Not only is there no one "best" approach, there are **six fundamental problems that exist regardless of the specifications which can make all of these methodologies unfair for evaluating providers and potentially problematic for efforts to improve the quality of patient care**.

### 1. Many Patients and the Spending on Their Care Are Not Assigned to Any Provider

In most attribution methodologies, a large number of patients are not assigned to any physician practice or other provider, and consequently, the spending associated with those patients is ignored in spending analyses. This can cause distortions in comparisons of spending between providers and it can also create perverse incentives for the providers:

- Patients who are not receiving adequate preventive care will be excluded, and providers who take on care of these patients can be financially penalized.
- The patients most in need of care coordination will be excluded, and providers who provide coordination to complex patients may be financially penalized.
- Providers can be financially penalized for keeping their patients healthy.

### 2. Providers Cannot Control All of the Services and Spending Assigned to Them

Even when a patient's spending is attributed to a physician or provider organization, that does not mean the physician or organization could have controlled or influenced all of the services that generated the spending. In fact, *most* of the spending that is attributed to physicians in typical methodologies results from services delivered by *other* physicians. Moreover, a physician can easily be assigned accountability for services a patient received before the physician first became involved in the patient's care. This creates a perverse incentive for a physician not to become involved with a patient who already incurred significant healthcare spending earlier in the year, even though these are the patients who may most need additional help.

### 3. Providers Are Not Attributed the Spending For Many Services They Provide

Not only are providers assigned spending that they cannot control, most attribution systems fail to assign physicians the majority of patients they did care for or the majority of services they delivered. Spending on preventable conditions such as hospital-acquired infections may be assigned to the physicians who treated the conditions rather than those who caused them. Moreover, spending by physicians who are delivering large numbers of services inappropriately or fraudulently may not be assigned to them, making it difficult to identify them and address these problems.

In addition, many measures described as "total spending" or "total cost" frequently exclude spending on prescription medications. As a result, spending totals for patients who use more drugs but use fewer other services will appear artificially low compared to others. In addition, since some types of drugs are paid for through the patient's medical insurance and other types of drugs are paid through prescription insurance, spending tabulations for physicians whose patients differ in the types of drugs they need will not be comparable if the spending under prescription insurance is not included.

### 4. Spending Measures Do Not Distinguish Appropriateness of Services

Under typical accountability systems, no distinction is made between recommended services and inappropriate services. A provider who does a better job of delivering recommended services could be measured as having higher spending than a provider who fails to deliver recommended services or a provider who delivers services that are less expensive but inappropriate for the patient. This could have the unintentional side effect of encouraging providers to stint on desirable care to patients in order to reduce the total amount of spending.

### 5. Risk Adjustment Systems Do Not Adequately Adjust for Patient Needs

The risk adjustment methods used in most accountability systems do not effectively separate differences in patient needs from differences in the way providers deliver care.

- Most risk adjustment systems are designed to predict *spending* on patient care, not adjust for differences in patient *needs*. This can reinforce inappropriate spending, penalize efforts to reduce underuse, and cause providers to focus spending reduction efforts on the wrong patients.
- Most risk adjustment systems use *historical* information on patient characteristics, not the most current information on health problems that affect the services patients need. This can penalize providers who care for patients with many acute healthcare problems.
- The *same* risk score can be assigned to patients who need very *different* kinds of services from physicians in different specialties. This can distort spending comparisons for physicians, particularly primary care physicians.
- Most risk adjustment systems only use information available in *claims data* that does not completely or accurately measure differences in patient health needs.
- Most risk adjustment systems give little or no consideration to factors *other than health status* that can affect patient needs. For example, patients who have functional limitations are more likely to have higher healthcare spending, but measures of functional limitations are not included in typical risk adjustment systems.

### 6. Inadequate Adjustments Are Made for Structural Differences in Costs

Providers in rural areas and poor communities incur higher costs to deliver the same services to their patients than do other providers. Most accountability systems make only limited adjustments for these differences, if any, which can penalize providers for factors outside of their control and potentially make it more difficult for patients to access the care the need.

## BETTER WAYS OF MEASURING AND ASSIGNING ACCOUNTABILITY FOR HEALTHCARE SPENDING

Clearly, better methods of measuring spending and assigning accountability are needed. An effective methodology needs five capabilities:

1. Identifying the services and spending that can be controlled or influenced by each provider;
2. Identifying which services represent opportunities for reducing spending without harming patients;
3. Determining which patients have greater needs for services;
4. Adjusting for structural differences in costs for different providers; and
5. Comparing providers based on both costs and outcomes of care.

### 1. Identifying the Services and Spending Providers Can Influence

The first step in more effectively identifying ways of reducing healthcare spending without rationing and identifying the healthcare providers best able to make the reductions is to divide spending into categories that reflect differing levels of provider control or influence over services. Five such categories are:

**Spending Category 1**: Services both *ordered and delivered* directly by the physician or other provider who is being measured.

**Spending Category 2:** Services delivered by *other* providers that are *integrally related* to services delivered by the provider being measured.

**Spending Category 3:** Services delivered by *other* providers that resulted from *orders* or *referrals* from the provider being measured, and services delivered by the provider being measured in response to orders from other providers.

**Spending Category 4:** Services delivered by other providers that were *related* to services delivered or ordered by the provider being measured.

**Spending Category 5:** All other services the patient received that are *unrelated* to services delivered or ordered by the provider being measured.

Collectively, these five categories add up to the total spending on all services a patient received. The services included in each category will differ for different providers, but in all cases, a provider will have greater influence over the lower-numbered categories than the higher-numbered categories, so this categorization will better identify which providers could actually reduce spending than current attribution methodologies which simply attribute the spending in all five categories to a single provider who happened to provide a certain proportion of the total spending. Moreover, under this approach, *every* provider will have *all* of the spending *they directly control* attributed to *them*, and *all* of the spending on *all* patients will be attributed to *at least one* provider.

## 2. Identifying Subsets of Spending That Can Be Reduced

The five spending categories better identify which providers have the ability to control or influence different aspects of spending, but they give only limited indications as to which aspects of spending could be *reduced*. To better identify opportunities to reduce spending without rationing, it is desirable to further disaggregate Spending Categories 1-4 into four subcategories:

**Subcategory (a):** Services required to meet quality standards.

**Subcategory (b):** Services that are potentially avoidable, e.g., services such as MRIs for lower back pain, cardiac stress tests, and Cesarean sections that may provide significant benefit to some patients but relatively little benefit to others and in some cases may result in harm to the patient that outweighs the benefit.

**Subcategory (c):** Services needed to address potentially preventable conditions, i.e., situations where the *health condition itself* could potentially have been prevented if additional or different services had been delivered *at an earlier point in time*.

**Subcategory (d):** All other services ("typical services"). Even if there is not enough evidence about appropriateness or preventability to classify them in the other three categories, variation among providers in the number and types of "typical" services they use for similar patients could indicate opportunities for savings and areas where research is needed to develop appropriate use criteria.

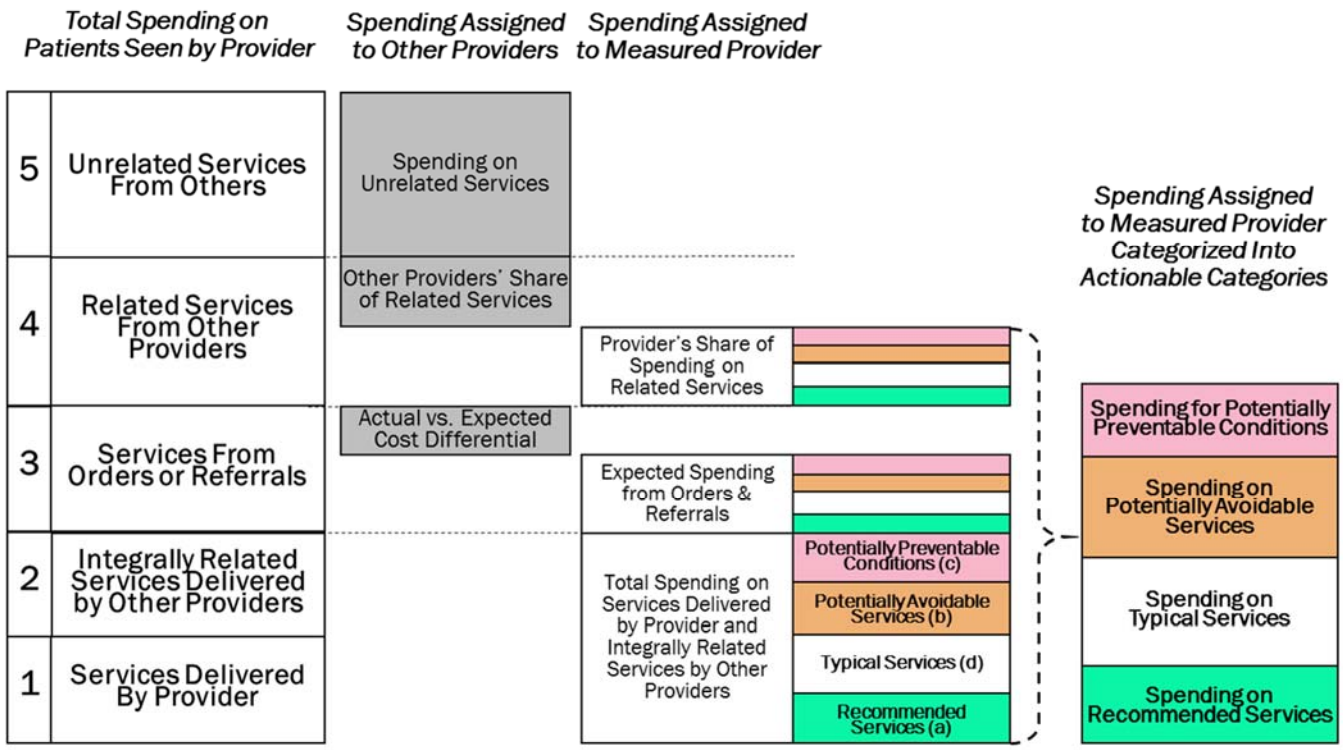
Figure A illustrates how the total spending on patients seen by a provider would be divided into the five spending categories based on the level of the provider's control and influence over the services delivered to the patient and then further divided into the four subcategories. For most providers, a significant share of the total spending for their patients will be assigned to other providers, but the spending that is assigned to them will be based on the services they can control or influence.

## 3. Adjusting for Differences in Patient Needs

Even with these better categorizations, comparisons among providers will not be meaningful unless they distinguish differences in spending that were associated with differences in the needs of the patients for whom the providers were providing care. The methods used to do this should address the many weaknesses of current risk adjustment systems in the following ways:

- **Disaggregating Spending into Subgroups of Patients with Similar Health Conditions.** Instead of using a *single* risk score to *adjust* spending, a better approach is to *compare* spending separately for *different subgroups* of patients, with each subgroup defined such that patients in that subgroup would be expected to need similar levels of services. Some risk adjustment systems have methods for grouping patients into clinically similar subgroups that can be used in this way for all types of spending.
- **Using Concurrent Risk Adjustment.** The patient categories should be based on complete information about the patients' health problems that occurred

FIGURE A
Identifying the Spending For Which a Healthcare Provider Can Be Accountable



during the time period in which spending is being measured, rather than only the kinds of historical information used in purely prospective risk adjustment systems.

- **Using Clinical Information from Electronic Health Records (EHRs) and Registries in Addition to Claims Data.** In many cases, the key information that distinguishes differences in patient needs is not captured at all in claims data, and so clinical data is also needed.
- **Disaggregating by Non-Health Factors to Identify Impacts on Spending.** Factors such as functional limitations can have a significant impact on the type and costs of healthcare services patients will need. *Disaggregating* spending into different categories of patients is also preferable to *adjusting* overall spending based on patient characteristics because it enables disparities between different groups to be measured and acted upon, rather than hidden inside a risk adjustment formula.

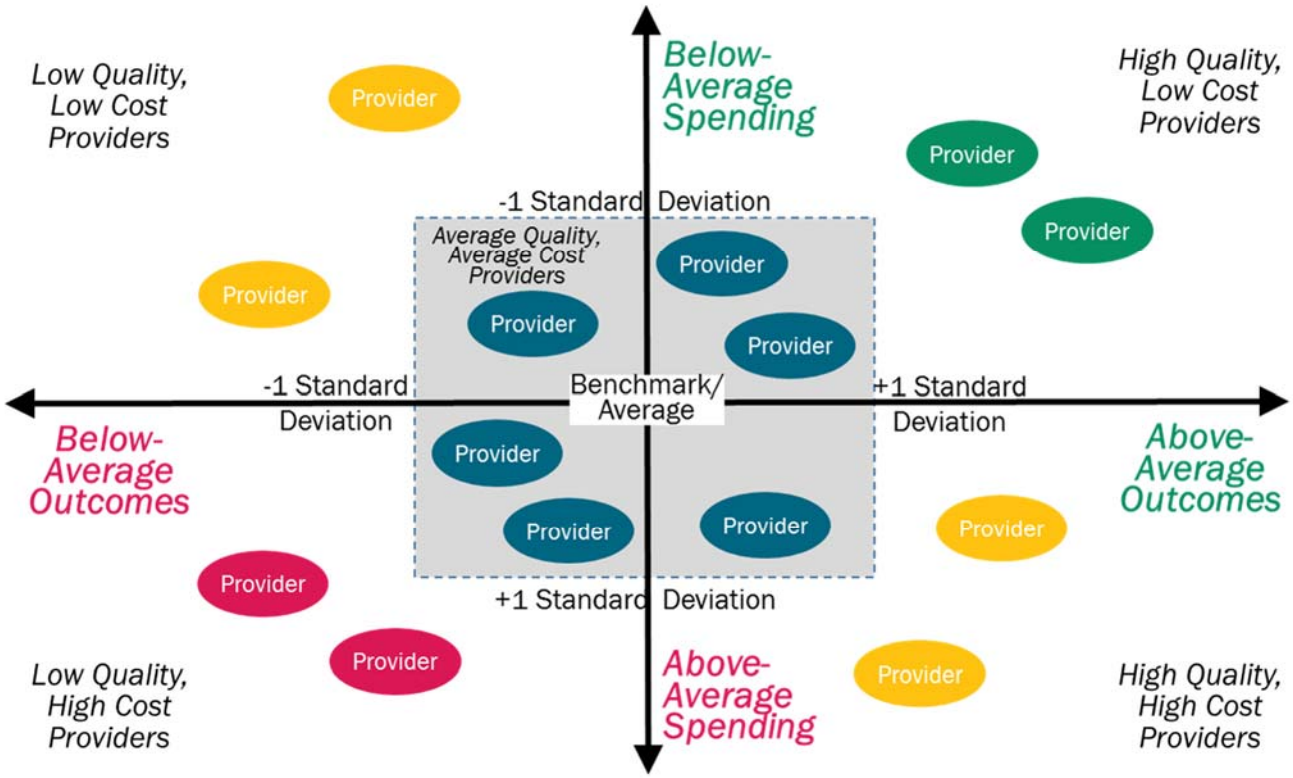
## 4. Comparing Providers That Are Comparable

Additional steps must be taken if comparisons are going to be made between providers who face structural differences in costs, e.g., comparing physician practices and hospitals in both urban and rural areas. One approach is to adjust the spending of each provider for their relative differences in costs, if these differences can be estimated; an alternative approach is to only compare spending between providers with similar structural characteristics (e.g., only comparing rural providers to other rural providers).

## 5. Assessing Differences in the Outcomes of Services and Spending

Since healthcare spending is not an end in itself, but a means to achieve better health and a higher quality of life for patients, it is important to distinguish physicians, hospitals, and other providers that spend more and achieve better health outcomes from those providers that spend more but do not achieve better outcomes and from those providers that spend less but achieve significantly poorer outcomes. This cannot be done through simplistic calculations such as dividing a measure of quality by a measure of spending. Since health outcomes and spending are measured on different scales and since different people may place different dollar values on the same outcome, the best approach is to show how providers differ on both measures, similar to the chart in Figure B, so that individual patients and payers can make their own judgments about which providers are "better." Because there will be small variations in quality and spending from year to year and provider to provider based on variations in patient needs and other factors which cannot be accurately measured, only providers whose performance is *significantly* better or worse than others should be identified as delivering higher or lower value care.

FIGURE B
Comparing Providers on Both Spending and Outcomes



## AN EXAMPLE OF MORE ACTIONABLE ANALYSIS

To see how spending analyses using the categories and subcategories defined above would differ from current methods, consider a hypothetical patient who receives the following health care services during the course of a year:

- In January, the patient visits his primary care physician (PCP) complaining of mild chest pain while exercising. The primary care physician orders a cardiac stress test to help determine if the patient is at risk of a heart attack.
- In February, the stress test is performed and the cardiologist who reviews the results determines there is no indication of significant coronary artery blockage. The cardiologist sees the patient in his office to explain the results, determines that the patient has risk factors for a future heart attack, and orders recommended medications.
- The patient has also been having lower back pain. The patient does not consult with the PCP about this problem, but contacts a neurosurgeon directly and schedules an appointment in March. The neurosurgeon evaluates the patient and recommends spinal surgery.
- In April, the neurosurgeon performs the surgery on the patient at a medical center fifty miles from where the patient lives. An anesthesiologist is also involved in the case.
- After discharge, the patient decides to go to a local skilled nursing facility for physical therapy and rehabilitation rather than for outpatient physical therapy, and the patient's insurance approves that service.
- The patient develops an infection at the site of the surgery and is admitted to a community hospital in May. A hospitalist successfully treats the infection and the patient is discharged.
- The hospitalist recommends that the patient see a primary care physician regularly in the future. The patient decides to use a different primary care physician than the one he had seen in January.
- The patient visits the new primary care physician in June and again in October, and the new PCP makes sure that the patient is up to date on all preventive care. The PCP finds that the patient has not had recommended screening for colon cancer and orders a colonoscopy.
- A gastroenterologist performs the colonoscopy in November and finds no evidence of cancer. The gastroenterologist performs the colonoscopy at the community hospital and uses an anesthesiologist to administer sedation.

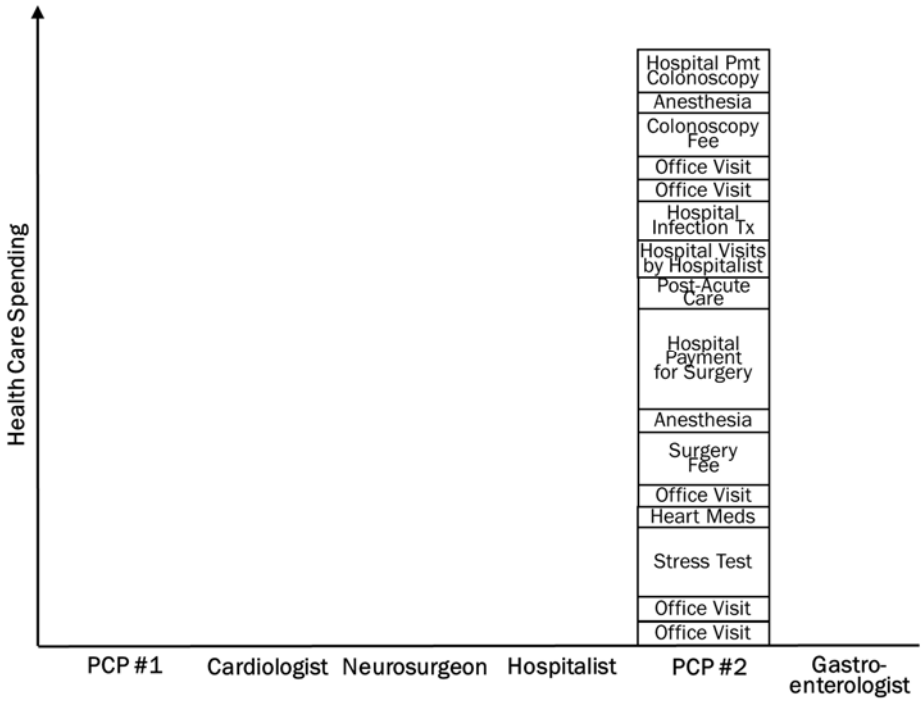
This patient has received services from a total of eight different physicians and two hospitals. As shown in Figure C, accountability systems that attribute spending based on primary care visits would attribute all of those services to the patient's new primary care physician (because the PCP had the largest number of visits with the patient), including the stress test, the back surgery, and the hospital readmission after the surgery, even though those services occurred before the new primary care physician had met the patient for the first time. None of the services would be assigned to the other physicians who actually delivered or ordered them.

As shown in Figure D, the spending category system described above would assign individual components of spending to each of the physicians who had control over each service, rather than assigning total spending to a single physician or to no physician at all:

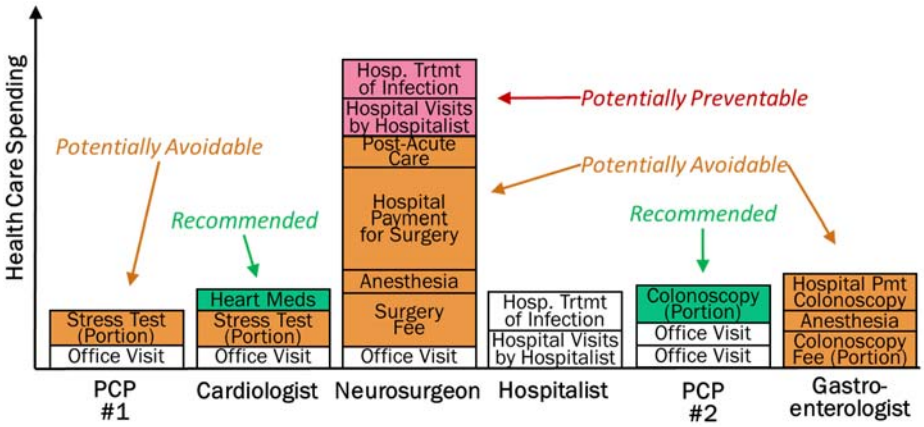
- The visit in January to the patient's first primary care physician would be placed in Spending Category 1(d) for that PCP. A portion of the cost of the stress test the PCP ordered would be included in Spending Category 3(b) for that PCP, since stress testing is often overused for low risk patients.
- The remaining portion of the cardiac testing would be included in Spending Category 3(b) for the cardiologist. The office visit with the cardiologist would be included in Spending Category 1(d) for the cardiologist, and the full cost of the heart medications ordered for the patient would be included in Spending Category 3(a) for the cardiologist, since they are recommended by guidelines.
- The neurosurgeon's fees would be included in Spending Category 1(b) for the neurosurgeon since spinal surgery for back pain is a frequently overused procedure. The spending for the anesthesiologist and the hospitalization would be included in Spending Category 2(b) for the neurosurgeon since they were integrally related services. The cost of the post-acute care would be included in Spending Category 4(b) for the neurosurgeon, since it was related to the procedure the physician performed in the hospital, even though the surgeon did not order inpatient rehabilitation or choose which facility was used.
- The admission to the community hospital was a complication of the surgery, so the hospital and hospitalist payments are including in Spending Category 4(c) for the neurosurgeon and the teaching hospital, as well as in Spending Category 1(d) for the community hospital and hospitalist.
- The two office visits with the second PCP are included in Spending Category 1(d) for that PCP. A portion of the cost of the colonoscopy ordered by the PCP is included in Spending Category 3(a) for the PCP.
- The remainder of the cost of the colonoscopy, including the cost of the anesthesiologist and the hospital, is included in Spending Category 3(b) for the gastroenterologist, who is responsible for the fact that the colonoscopy is performed in a hospital using an anesthesiologist, making it more expensive than if it were done in the gastroenterologist's office using an alternative form of sedation.

This approach shows which of the physicians and hospitals are in the best position to reduce potentially avoidable services and potentially preventable conditions, and it enables comparisons of each of the physicians to peers based on the services they are able to control or influence. In this example, it is clear that the potentially avoidable spinal surgery had the single biggest impact on total spending for the patient during the course of the year, and it also involved the largest amount of potentially preventable spending as a result of the infection that required the patient to be hospitalized a second time. The neurosurgeon was responsible for all of this spending; neither of the patient's primary care physicians had any involvement in the patient's decision to see the neurosurgeon or in the neurosurgeon's decision to perform spinal surgery. Other, smaller opportunities to reduce spending include the potentially avoidable stress test ordered by the first PCP, and the appropriate but more-expensive-than-necessary colonoscopy delivered by the gastroenterologist.

## FIGURE C
## Attributing Total Spending to the Primary Care Physician with the Plurality of Visits




## FIGURE D
## Assigning Components of Spending to the Providers Best Able to Control Them



## MOVING FROM MEASUREMENT TO ACCOUNTABLE PAYMENT

Developing more actionable information on health care spending is a critical first step in trying to reduce or control health care spending. However, fee-for-service payment systems create significant barriers to reducing spending without harming patients, such as failing to pay or paying inadequately for services that could lower overall spending, and financially penalizing physicians and hospitals for reducing unnecessary services and improving quality.

These barriers cannot be solved by merely adding bonuses or penalties based on health care spending measures on top of the *current* fee-for-service payment system. Moreover, if the measures used for the bonuses and penalties are flawed, they can create perverse incentives for providers to avoid caring for patients who could benefit the most from improved care.

Instead, *different* payment systems are needed to truly overcome the barriers. True payment reforms – bundled payments, warrantied payments, episode payments for a procedure, condition-based payments, and global payments – give physicians, hospitals, and other providers the *flexibility* to redesign care in more efficient and effective ways, but also the *accountability* for ensuring that the care is delivered in the highest quality, most affordable way.

The structure for aggregating and disaggregating data described earlier provides data in exactly the format that is needed for defining and pricing better payment systems:

- Distinguishing between services delivered or ordered by a provider versus those delivered or ordered by other providers enables payment systems to give providers accountability for aspects of spending they can control or influence as well as the flexibility to change the way services are delivered in order to impact spending, while avoiding penalizing providers for services and spending they cannot control.
- Accurately determining which types of patients will need more services ensures that payment systems will pay adequately for patients with more serious health problems, and it will encourage providers to take on the care of such patients, rather than discourage them from doing so.
- Identifying areas of overuse and underuse in spending enables prices to be set at levels that are financially feasible for providers while avoiding creating pressure to stint on needed care.

## OBTAINING THE DATA NECESSARY FOR IMPLEMENTATION

In order to develop better spending measures and better payment systems, investments will need to be made in effective data collection and analytics, including:

- All-payer claims data on the services patients are receiving and the amounts being spent on those services;
- Clinical data on patients, particularly data from patient registries; and
- Data on patient outcomes.

A growing number of states and communities are working to assemble these types of data and use them to create information on the quality and cost of care that can help providers, purchasers, and patients to redesign healthcare delivery and payment systems. These state and local efforts are more likely to be successful than a one-size-fits-all national approach. Multi-stakeholder Regional Health Improvement Collaboratives can represent the most efficient and effective way to collect and analyze information in a way that all stakeholders can trust. However, all stakeholders will need to contribute sufficient funding to ensure there are adequate resources available to support this important but complex work.

## CONCLUSION

Controlling healthcare spending without harming patients will depend on active engagement and strong leadership from physicians, hospitals, and other providers. Poorly designed measurement, attribution, and accountability systems not only fail to provide the actionable information providers need, they can discourage providers from making feasible changes by demanding they control services and spending that are beyond their range of influence. Pay for performance and shared savings programs based on spending measures not only fail to resolve the barriers to change created by fee-for-service payment, they can further discourage action by penalizing providers based on flawed systems of assigning accountability.

Fortunately, there are better ways to analyze spending that can help physicians, hospitals, and other providers identify opportunities to achieve better outcomes at lower costs. There are also better ways to pay providers that will enable them to redesign care to implement those opportunities in ways that are financially feasible for them. Although significant investments of time and money will be needed to create better analyses of spending and to design and implement better payment systems, the potential savings from reducing healthcare spending provide the opportunity for a significant return on investment for all stakeholders. Many states and regions already have laid the foundation for this through initiatives to assemble data, create analytic capability, pursue innovative payment reforms, and facilitate collaboration among payers, providers, and other stakeholders. These communities can lead the way for the rest of the country if they receive the necessary support to do so.

# I. THE PURSUIT OF ACCOUNTABILITY FOR HEALTHCARE SPENDING

For over a decade, there has been a wide array of efforts designed to improve the *quality* of healthcare services in the United States. Hundreds of quality measures have been developed, data on many of these measures have been collected by both payers and providers[1], a number of the measures have been publicly reported, and various pay-for-performance systems using these measures have been created to reward healthcare providers for improving quality, to penalize them for poor quality, or both.

More recently, national concern about the high cost of health care services has led to a parallel set of efforts focused on health care *spending*. Various measures of spending[2] and resource use are being developed and data on these measures are being collected and reported in various formats. Increasingly, payment systems are being modified to use these types of measures to reward or penalize healthcare providers in order to encourage lower and slower-growing spending.

## A. Measurement and Reporting on Spending

In an effort to encourage and assist healthcare providers to control health care spending, Medicare and many commercial health plans have been giving providers information and analyses about the services their patients have received and the spending associated with those services. For example, for several years, the Centers for Medicare and Medicaid Services (CMS) has been giving physicians Quality and Resource Use Reports (QRURs), which show measures of both the quality of care and spending for patients to whom the physicians provided services.[3] The measures of spending not only include services delivered by the physicians themselves, but services delivered by all other physicians, hospitals, and other providers. In several communities, multi-stakeholder Regional Health Improvement Collaboratives are also producing analyses of spending to help physicians, hospitals, employers, and health plans identify the causes of growth in healthcare spending and the opportunities to reduce spending.

Although the reports given to providers are generally not available to the public, some measures of spending are being publicly reported. For example, on the federal Hospital Compare website[4], the public can see the Medicare Spending Per Beneficiary measure for individual hospitals. This is a measure of the total spending on all services that Medicare beneficiaries received in the 3 days prior to a hospital admission and in the 30 days after discharge from the hospital, including services from physicians, post-acute care providers, and others.[5]

## B. Modifying Provider Payment Based on Spending

A growing number of payers are now using spending measures not just to *inform* efforts to reduce spending, but to *assign accountability* to individual providers for spending and to reward or penalize them on that basis. There are three basic ways in which this is done:

- **Pay-for-Performance.** In pay-for-performance systems, the amounts that providers are paid for the services they deliver are increased or decreased by a pre-defined amount if one or more measures of spending on healthcare services for their patients is better or worse than a benchmark. For example, the Centers for Medicare and Medicaid Services (CMS) will be adjusting Medicare payments to individual physicians and physician practices based on measures of "Total Per Capita Cost" and "Medicare Spending Per Beneficiary," and CMS will be adjusting Medicare payments to hospitals based on the Medicare Spending Per Beneficiary measure.[6]
- **Shared Savings/Shared Risk.** Shared savings and shared risk programs are a variation of pay-for-performance in which a provider receives a bonus payment or penalty that is proportional to the amount by which total spending for their patients was under or over a benchmark spending level. For example, CMS uses a shared savings methodology for primary care practices in its Comprehensive Primary Care Initiative demonstration program[7] and for groups of providers designated as Accountable Care Organizations in the Medicare Shared Savings Program[8] and the Pioneer ACO Program[9]. In the extreme case, where the provider is responsible for 100% of the difference between actual spending and the benchmark, the provider is, in effect, being given a virtual budget for all of the care its patients need.
- **Tiering and Narrow Networks.** Instead of directly changing the amount of payment the provider receives for individual services, *tiering* gives patients incentives and/or disincentives to use providers based on how those providers rank on spending measures. For example, a patient may have to pay more to obtain a service from a provider rated as "more expensive" on a total spending measure even if the price of that particular service is lower than what is charged by providers rated as "less expensive" overall.[10] *Narrow networks* may require patients to use only providers that are measured as "less expensive" based on the spending measures. A growing number of commercial health plans are using these approaches.

## C. The Need for Better Approaches to Measurement and Payment

There are a variety of methodologies that can be used for measuring and analyzing spending. Any methodology has strengths and weaknesses, and the most appropriate approach will depend on the way healthcare services are organized in the community and the types of services and spending being examined. When measurement is used to identify potential opportunities for reducing spending and is then followed by more detailed analysis, weaknesses in the methodology can be identified and corrected to more effectively support efforts to reduce healthcare spending.

However, when spending measures are used to modify payments to providers or to steer patients to particular providers, there is typically not a way to identify and correct weaknesses in the measurement methodology or to customize the analysis to the unique characteristics of individual communities. Unfortunately, the methodologies typically being used by payers to modify payments to providers have significant problems that can make them unfair as a way of evaluating individual providers and potentially problematic for efforts to improve the quality of patient care. Section III of this report describes in detail how:

- Many types of patients and spending are excluded from the measures used for accountability, including patients who may represent significant opportunities for reducing spending through improvements in care.
- Many providers are being held accountable for aspects of spending they do not have the ability to control.
- Some providers are not being held accountable at all for the spending for which they are responsible.
- Providers can be inappropriately penalized for caring for sicker patients and for providing the services needed to achieve better health outcomes for their patients.
- Spending measures are not structured and analyzed in the most effective ways to help providers identify ways to reduce spending without harming patients.

Although using these flawed measures to modify payments is being portrayed by many payers as a move toward "value-based payment," the problems with the measures can actually penalize providers financially when they redesign care in ways that will lower spending, and they can create perverse incentives for providers to avoid caring for many of the patients who could benefit the most from improved care.

Fortunately, there are better ways to analyze spending, assign accountability, and pay providers that can mitigate or solve these problems:

- Section IV of this report describes better ways to measure the services and spending that providers can influence, to identify the subsets of spending that can be reduced without harming patients, to adjust measures of spending based on patient needs, and to compare providers in fair and effective ways.
- Section V illustrates how spending analyses constructed in this way can help communities and providers to reduce spending without harming patients.
- Section VI explains how appropriately designed spending analyses can facilitate the design and implementation of payment systems that can support better care for patients at lower costs while enabling health care providers to remain financially viable.
- Section VII describes the types of data needed to support better measurement systems and payment reforms, and also describes how those data can be obtained and used most efficiently and effectively by communities seeking to improve the quality and affordability of their health care systems.

# II. CURRENT METHODS OF ASSIGNING ACCOUNTABILITY FOR SPENDING

The methodologies used by the Centers for Medicare and Medicaid Services (CMS) and most commercial payers to assign accountability for spending to providers in both reporting and payment systems have three basic components:

- The methodologies first define the subsets of healthcare spending that will be measured;
- The methodologies then attribute these subsets of spending for each individual patient to one or more providers; and
- Finally, the methodologies adjust the attributed spending in various ways in order to allow comparisons of spending across different providers and different communities.

## A. Attributing Spending to Providers

### Attributing Total Spending for a Patient's Care

To date, the approach most frequently used by Medicare and other payers has been to assign accountability to a single physician, hospital, or other provider for *all* of the spending on *all* of the healthcare services received by a patient during a particular period of time, regardless of which physicians, hospitals, or other providers actually delivered those services. For example:

- In the Hospital Value Based Purchasing Program established by CMS, hospitals are evaluated using the Medicare Spending Per Beneficiary measure for any Medicare patient admitted to that hospital. The measure includes the spending on any service delivered to the patient by any provider during the three days prior to an admission to the hospital and during the thirty days after discharge from the hospital, including services unrelated to the condition for which the patient was hospitalized.
- In the Physician Value Based Payment Modifier established by CMS, physicians are evaluated using the Total Per Capita Cost measure for a patient assigned to them. The measure includes the spending on any service delivered to the patient by any provider during the course of the year.

In general, only the healthiest patients will receive all of their healthcare services from a single physician[11] in his or her own office, and the small number of services most of these patients receive represent a small share of total health care spending, at least in the short run.[12] Most healthcare spending is associated with patients who received services from two or more physicians and from other healthcare providers such as a laboratory or a hospital[13], and in general, those physicians and providers will not all be part of the same healthcare organization.[14] Similarly, many patients who are hospitalized will receive significant "post-acute care" services from physicians, home health agencies, skilled nursing facilities, etc. in the weeks after discharge, and in many cases, those providers will not be owned or operated by the hospital or by the organization that operates the hospital. Moreover, most patients have the type of health insurance that does not require the patient to obtain approval from one provider before receiving services from other providers.[15] This means that for most patients, no single physician, hospital, or other provider organization is *prospectively* given responsibility for determining which healthcare services a patient will receive during any extended period of time or how much will be spent on those services.

Consequently, in the accountability systems used by Medicare and other payers, statistical rules are used to *retrospectively* assign responsibility to an individual physician, physician practice, hospital, or other provider for all of the services that a patient received over a specific period of time. This is typically described as "attributing" the spending to that particular provider.

The most common approaches to attribution are:

- **Attributing total spending to a primary care physician.** In this approach, all of the primary care services[16] that the patient received during a "lookback period" (e.g., the prior calendar year or the most recent 12 months) from any primary care physician[17] are identified. If the patient received the majority or plurality[18] of primary care services from a particular primary care physician compared to other primary care physicians, the patient is "assigned" to that primary care physician. The total spending on *all* services that patient received during the prior year – not just primary care, but services from all other specialists, hospitals, etc. – is then attributed to that primary care physician.
- **Attributing total spending to a physician (primary care or specialist).** In this approach, all of the office visits or physician services that the patient received over the past year are identified (not just primary care services). The patient is then assigned to the physician who provided more office visits or who billed for more services to that patient than any other physician, regardless of the physician's specialty. All of the spending associated with the patient is then attributed to that physician, including spending for services provided by other physicians, hospitals, etc.
- **Attributing total spending to a multi-provider organization.** A combination of the above rules may be used for organizations involving both primary care physicians and specialists. For example, in the Medicare Shared Savings Program, spending for Medicare ben-

eficiaries is attributed to an Accountable Care Organization (ACO) by first attempting to assign the beneficiary to a primary care physician (PCP) based on which PCP delivered the plurality of primary care services to that beneficiary; if the beneficiary received no primary care services from primary care physicians, then an attempt is made to assign the beneficiary to a specialist based on primary care services provided by a specialist. If the beneficiary is assigned to either a primary care physician or specialist affiliated with the ACO, then the ACO is assigned all of the spending for that beneficiary.[19]

- **Attributing total spending to all involved providers.** A less common approach to attribution does not attempt to select any *single* physician or organization, but attributes a patient (and the total spending on that patient) to *every* physician or provider who delivered a predefined *minimum proportion* of the total care during the lookback period (e.g., any physician who had more than 30% of the patient's total physician visits during the prior year).

There are many variations on these approaches. As shown in Figure 1, the formulas for performing attribution have multiple components, and for each component, there are several choices that can be made about how the attribution is done.

The results of attribution differ significantly depending on which of these options are chosen. For example, a 2010 study by Ateev Mehrotra and colleagues compared the results of applying six different attribution methodologies to total spending for the same group of commercially insured patients. They found that the percentage of patients who would be assigned to a single physician ranged from a low of 20% to a high of 51% depending on which attribution rule was used, and the percentage ranged between 42% and 69% if total spending was assigned to any physician who had more than 30% of the visits or spending for a patient.[20]

## FIGURE 1
### Components and Alternatives Used in Attribution Methodologies

| METHODOLOGY COMPONENT | EXAMPLES OF DIFFERENT SPECIFICATIONS |
|---|---|
| Measure Used for Attribution | • Number of Evaluation & Management visits<br>• Number of visits or services of any type<br>• Amount of payments for services<br>• Provision of a specific type of service |
| Providers Eligible for Attribution | • Primary care physicians<br>• Primary care providers (including physician extenders)<br>• Specialists providing primary care services<br>• Physicians providing any services |
| Magnitude of Measure Needed for Attribution | • Plurality within the measurement period<br>• Majority within the measurement period<br>• Specific threshold (e.g., 30% or more of visits)<br>• No threshold (i.e., any number of services) |
| Tie-Breaker (secondary measure used to assign accountability if two or more providers are equal on the primary measure) | • Most recent Evaluation & Management visit<br>• Most recent service<br>• Plurality/majority of services in more recent subset of time |
| Time period for measurement ("look back period") | • 1 year<br>• 18 months<br>• 2 years |
| Frequency of attribution | • Monthly<br>• Quarterly<br>• Annual |

With so many different parameters, and no "right" way to define any of them, it is not surprising that the details of the attribution methodologies differ from payer to payer. As a result, the spending attributed to a provider by one payer might be very different than the amount attributed to that provider by another payer, even if the patients insured by each payer had similar characteristics and received similar services.

### *Defining "Total" Spending*

The term "total" spending implies that all types of spending are included. However, as a practical matter, accountability measures can only be based on spending data that are accessible to the entity that is producing the measures and that are uniformly available on all patients. Spending measures calculated by payers are generally limited to spending on the services they pay for, and so they may exclude spending on large categories of services that are financed by other payers, such as prescription drugs. For example, the "total per capita cost" and "total spending per beneficiary" measures used by CMS in its value-based purchasing programs do not include information on spending on prescription drugs obtained through outpatient pharmacies, partly because these drugs are paid for by health plans under a separate Medicare program (Part D), and partly because not every Medicare beneficiary has prescription drug coverage through Part D. In addition, "total" spending measures generally do not include spending by patients on items that are not covered by insurance at all, such as over-the-counter medications.

## Defining and Attributing Episodes of Care

Instead of trying to assign accountability for *total* spending to a physician or provider organization, an alternative approach is to assign accountability for a *subset* of total spending called an "episode." An episode is a group of services a patient receives during a particular period of time and that are clinically related to each other in some fashion. For example, if a patient was hospitalized for repair of a broken hip at one point during the year and received a cardiac stent at a later point in the year, the services related to the hip surgery would be grouped into one episode and the services related to the cardiac stent would be grouped into another episode. Each episode would then be assigned separately to a physician, hospital, or provider organization based on the same kinds of rules described earlier, e.g., the physician who had the largest share of services or billings during the hip surgery episode would be assigned accountability for that episode, and the physician who had the largest share of services and billing during the cardiac stent episode would be assigned accountability for that episode.

For example, as part of the Medicare Physician Value Based Modifier program, the total spending for an episode of care surrounding a patient's hospitalization (as determined by the Medicare Spending Per Beneficiary measure) will be assigned to the physician group that furnished the plurality of Medicare Part B services during the hospitalization portion of the episode.

The same types of methodological choices described in Table 1 are involved in assigning an episode to a provider. In addition, however, choices must be made about the definition of the episode itself, i.e., which services will be included in a particular episode and which will be excluded. For example, if a patient has hip replacement surgery and later has to return to the hospital for a problem with their hip, the re-hospitalization might be included in the same episode as the hip surgery if the readmission occurred within 30 days after discharge from the surgery, but it might be treated as a new episode if it occurred farther in the future. Moreover, the re-hospitalization might be excluded from the surgery episode if the hip problem was determined to be unrelated to the surgery (for example, if the patient was in a car accident and happened to break the same hip that had just been operated on). If a patient is treated for two different problems during a hospitalization (e.g., if a patient has a heart attack, falls and breaks a hip, and has both hip surgery and an angioplasty during the hospitalization), and if the hospital is paid a single amount for the entire stay, rules must be established for how to separate the spending for the hospital services into separate episodes for the hip surgery and for the angioplasty.

For patients with chronic diseases, the concept of "episode" has been expanded to include all of the care that occurs for the chronic disease during a 12 month period, since, by definition, chronic diseases are never cured and therefore not time-limited. However, unlike with acute episodes, the end of the 12 months does not actually signal a change in the patient's condition, it is merely an arbitrary dividing point.[21] As a result, the differences in spending in two consecutive 12 month "episodes" for a patient will depend heavily on whether a high-cost service or acute exacerbation of the disease occurred before or after the arbitrary end of the 12 month period. Under typical attribution rules, a different physician may be assigned the spending in each year depending on which physician happened to provide more services during that inherently arbitrary period.

A number of different episode "groupers" have been developed to make these determinations.[22] These computer programs make certain assumptions about which services should be grouped together using information derived from claims data about which diagnoses were associated with the services and when the services occurred. 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.[23] 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.[24] 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.[25]

This means that whether the episode is assigned to a provider will first depend on which episode grouper is used, since different groupers may assign different services (and the providers who delivered them) to different episodes. The assignment will then also depend on which attribution rule is used to assign the episodes to providers. For example, the Mehrotra study cited earlier found that approximately one-half of physicians would be assigned to a different cost category (high, average, or low) depending on which of several rules was used for attributing episodes to physicians.[26]

## B. Determining Whether Spending is "High" or "Low"

Ideally, after determining how much was *actually* spent on healthcare services for a provider's patients, one would determine which services would have been *appropriate* for the provider's patients based on medical evidence and the most *efficient cost* of delivering those appropriate services. Then the actual and appropriate/efficient spending levels would be compared to determine whether the provider's spending was "high" or "low."

However, despite a wide of range of efforts to develop evidence about the appropriateness and effectiveness of healthcare services, in most cases it is not possible to define with any accuracy what services are most appropriate to address the unique needs of individual patients. There is also relatively little information available to indicate how much it would cost for providers to deliver services if they were operating in the most efficient manner possible.

Consequently, the method typically used to determine whether a provider's spending is "high" or "low" is to compare spending on that provider's patients to the spending on the patients of other providers.[27] In order to make "apples to apples" comparisons, two types of adjustments are typically made to the spending measures:

- **Risk Adjustment:** Since sicker patients will need more services, there is general agreement that comparing different providers on spending for different groups of patients requires a method of adjusting for differences in the needs of those patients.
- **Price Adjustment:** There are also differences in what different providers are paid for what is ostensibly the same service for similar patients. In some cases, these are intended to reflect legitimate structural differences in those providers' costs, and adjustments need to be made for these structural differences in order to make fair comparisons of spending for different providers and for providers in different communities. In other cases, the actual payment amounts are not known, and so estimates of the differences in costs/prices between different services must be used in order to compare spending.

### Risk Adjustment

The most common approach used to adjust for differences in patient characteristics is to calculate a "risk score" for each patient based on the extent to which that patient has characteristics that are associated with higher spending. Then the total spending on a provider's patients is divided by the average risk score for those patients in order to compute a ratio called the provider's "risk adjusted spending." If a provider's risk-adjusted spending is higher than the risk-adjusted spending of other providers, it is presumed that the provider is delivering services inefficiently or inappropriately to the patients.

> ***EXAMPLE:*** *The Medicare program uses the Hierarchical Condition Category (HCC) system[28] to take information on the health issues an individual patient had during the preceding year in order to calculate a single numeric "risk score" for that patient. These risk scores are averaged across all of the patients associated with a provider or health plan to determine the Risk Adjustment Factor (RAF) for that group of patients. Then the total spending during the year for a group of patients is divided by their RAF score to determine the risk-adjusted spending.*

As shown in Figure 2, a provider that has higher unadjusted per patient spending than others may actually have lower spending on a risk-adjusted basis if the risk scores for that provider's patients are higher than those for other providers.

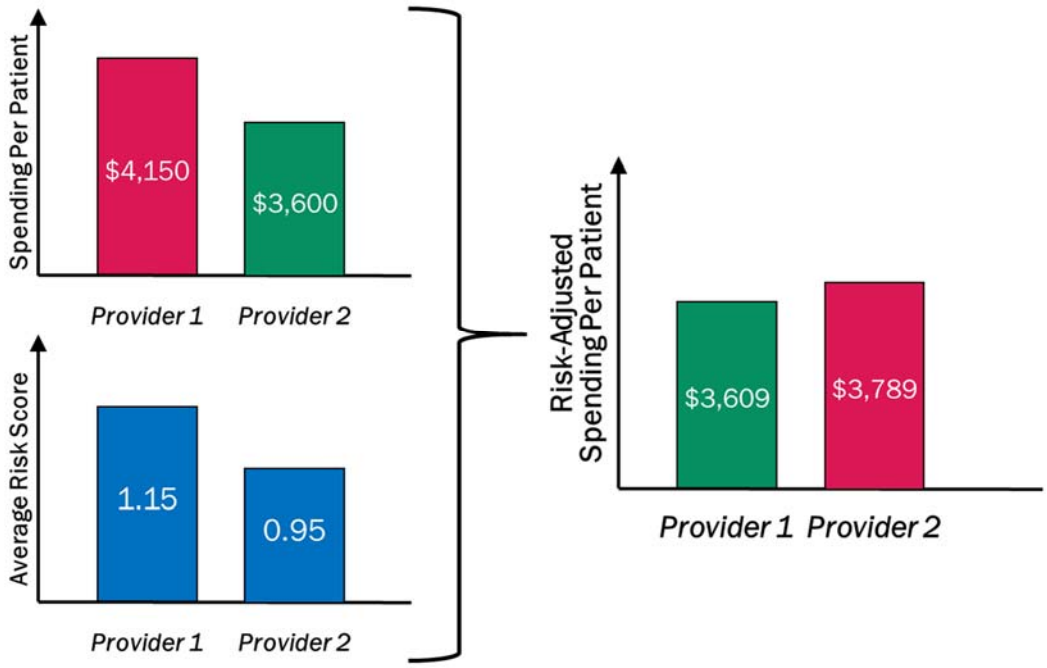
Risk adjustment introduces an additional set of methodological choices into the overall process of measuring and assigning accountability for spending. First, decisions must be made as to which characteristics of the patient will be considered in determining that patient's risk score. Then, a methodology is needed for assigning relative weights to those characteristics based on the presumed impact they have on a patient's need for services, and for combining those weights into a single risk score. There are multiple risk adjustment systems available, and each uses a different methodology for what patient characteristics are considered and how those characteristics are weighted and combined to determine a risk score.[29]

### Price Adjustment

As a matter of policy, Medicare recognizes that there are certain structural factors that cause different providers to incur different costs in delivering the same service. For example, Medicare pays physicians and hospitals higher amounts if they are located in communities with higher costs-of-living, and Medicare pays teaching hospitals higher amounts than non-teaching hospitals. In order to avoid having the higher-paid providers inappropriately classified as "higher spending" for delivering the same services because of the higher payment amounts Medicare has established for them, Medicare analyses of spending are generally based on "price-standardized" spending amounts. Price standardization is performed by calculating the ratio of the payment made to each provider for a particular service and the amount that would be paid for that same service to a reference provider. Actual spending by a provider is then adjusted upward or downward by this ratio to calculate the price-standardized spending.[30]

Anyone can make these adjustments for Medicare spending because the amounts that the Medicare program pays for each service from individual providers is public knowledge, but adjustments are more challenging for commercial spending because the amounts that commercial health plans pay providers for services are generally not public knowledge. In fact, because payment levels are generally not known, in order to compare *spending* between providers rather than merely service *utilization*, data on the number and types of services delivered by providers must be adjusted in some way to reflect the differences in costs or payments between services. One approach to doing this is to use Medicare payment rates for the same services in place of the unknown commercial payment rates. Another approach is to use the average of the amounts paid by multiple commercial payers (although this requires knowing the amounts each commercial payer pays in order to calculate the average, it does not require making the payer-specific amounts public). Whatever method is used, the payment levels then need to be stand-standardized across different types of providers in order to create fair comparisons of spending.

FIGURE 2
Using Risk Scores to Adjust Spending Comparisons



# III. PROBLEMS WITH CURRENT METHODS OF ASSIGNING ACCOUNTABILITY

Although most payers use *similar* methodologies for attribution, episode definition, and risk adjustment, the details of the methodologies differ from payer to payer. As noted earlier, comparisons have shown that the resulting spending measures can differ significantly depending on those details. If different payers use different specifications, a provider could receive a different rating on spending from each payer even if the provider was delivering care in the same way to each payer's patients.

Moreover, the more health plans and other payers that are involved with an individual provider's patients, the smaller the number of patients on which each payer's spending measure for that provider will be based, and the less reliable each measure will be. Differences in methodologies and small sample sizes reduce the credibility and utility of spending measures generated by individual payers, particularly in communities where there are many different health plans.

Consequently, if the goal is to help providers find ways to reduce spending, measures generated using all-payer data with a single methodology will provide more reliable and actionable information. Regional Health Improvement Collaboratives in several communities are beginning to produce spending measures in this way in order to help identify opportunities for controlling healthcare spending as well as improving quality.[31]

Using a single methodology requires deciding what that methodology should be. There is no one "best" approach to the specifications in the most commonly used methodologies; there are advantages and disadvantages to each depending on the goal of the measures (e.g., whether they are being used to help providers identify opportunities to reduce spending or to adjust their payments based on utilization of services) and the characteristics of the community where the measures are being produced (e.g., whether patients change health plans frequently, whether most physicians are part of large multi-specialty groups or single-specialty practices, etc.).

More importantly, **there are six fundamental problems with most current methodologies that exist under *all* of the specifications commonly used**. These problems can make the methodologies less effective in supporting efforts to reduce health care spending, unfair as a way of evaluating individual providers, and potentially problematic for efforts to improve the quality of patient care:

- Many patients and many aspects of spending are not attributed to any physician or other provider.
- The physician or organization that is attributed spending for a patient may not have been able to control or influence many of the services that patient received.
- Many providers are not attributed the spending that they can control.
- No distinctions are made between necessary and avoidable services.
- Comparisons of spending across providers do not adequately adjust for differences in patient needs.
- Comparisons of spending do not adequately adjust for structural differences in costs among providers.

As described in more detail in Section III-G below, the significance of each of these problems will vary depending on the community where spending measures are being created, how the measures are created, and the purposes for which the measures are being used. As described in Section IV, there are better ways to measure spending that can reduce or avoid these problems altogether. However, understanding the value of alternative methods requires an understanding of the problems with current methodologies, and these problems are described in detail in Sections A-F below.

## A. Many Patients and Their Spending Are Not Assigned to Any Provider

In most attribution methodologies, a large number of patients are not assigned to any physician practice or other provider. This is because many patients do not receive healthcare services from primary care physicians during the relevant measurement period and/or because they receive care from so many different physicians that no one of them provides a minimum percentage of the patient's care.

For example, when Hoangmai Pham and colleagues used an attribution rule based on the plurality of visits to primary care physicians in order to assign Medicare patients to physicians, they found that only 79% of patients could be assigned to a primary care physician. This was because 15% of the beneficiaries had visits only with specialists and 6% had no evaluation and management visits with any physicians at all.[32]

The number of unassigned patients is even higher for younger patients with commercial insurance. In the study by Mehrotra and colleagues, the highest percentage of patients for whom total spending could be assigned to a single physician (of any specialty) under either plurality or majority assignment rules was 51%, and only 69% could be assigned if the threshold was reduced to include any physician with more than 30% of visits or costs.[33] In a study conducted by HealthPartners in Minnesota, only 53.1% of members were attributed to a primary care physician using a rule based on the physician who delivered the plurality of evaluation and management visits, and no more than 68% were attributed under any of the rules tested.[34]

Even when more narrowly-defined episodes of care are assigned to physicians rather than all of the services and spending for a patient, a large number of episodes are unassigned. In the Mehrotra study, only about half (51-55%) of episodes could be assigned to a single physician if rules based on the plurality or majority of visits or costs were used, and only slightly more (54-58%) could be assigned if episodes were assigned to any physician with more than 30% of the visits or costs in an episode.[35] A study conducted by the Medicare Payment Advisory Commission found that a much higher percentage – 90% – of episodes could be assigned to a single physician using a 30% threshold (i.e., the episode is assigned to the physician with the largest share of services or spending, as long as that share is greater than 30%), but only 75% could be assigned using a 50% threshold. Not surprisingly, the percentage varied significantly depending on the complexity of the condition and the likely involvement of multiple physicians – with the 50% attribution threshold, only 53% of cerebrovascular episodes could be assigned to physicians whereas 88% of sinusitis episodes could be assigned.[36]

*Current attribution methodologies could encourage healthcare providers to avoid becoming involved in care of patients who use emergency rooms frequently or have complex problems requiring expensive healthcare services.*

There are at least three reasons why the large number of unassigned patients and unassigned services/spending is a problem:

- **Patients who are not receiving adequate primary care services will be excluded, and providers who take on these patients may be penalized.** It seems quite likely that many of the patients who are unassigned are those for whom the opportunities to reduce spending are the highest. For example, patients who visit the emergency room frequently and never see a primary care physician will not be attributed to any primary care physician or other provider under attribution rules that assign patients based on visits to primary care providers, and as a result, none of the services for these patients will be included in the spending measures. Similarly, the patient who avoids preventive care and then develops a serious health problem requiring hospitalization will not be assigned to a primary care physician or even a specialist, and again, this aspect of spending will not be measured. For example, in the HealthPartners study cited earlier, over one-fifth (22.5%) of all emergency room visits were made by patients who were not attributed to any primary care physician using a rule that attributed spending to the physician with the plurality of evaluation and management visits, and over one-sixth (17.3%) of all hospital admissions were for patients not attributed to any primary care physician.[37]

  This does not mean that all of these patients *should* be assigned to a primary care physician or any other physician. If no physician was actually managing these patients' care, it would be inappropriate to hold any physician accountable for the fact that the patient needed services that could have been avoided had they received better care. However, if the real goal of a payer or community is to reduce healthcare spending, rather than simply to measure physician performance, then the failure to identify and focus on these patients is a serious weakness. In most accountability systems, the spending on these patients is simply ignored, and measures and payments are based only on the patients who are assigned to providers.

  Moreover, the failure to include these patients creates a perverse penalty for providers who do become involved in the patients' care. For example, if a primary care physician begins seeing the patient who had been using the emergency room frequently, all of the patient's emergency room visits for the past year will be attributed to that primary care physician under most attribution systems, even though those visits occurred before the PCP first became involved in the patient's care. That will tend to increase the per-patient spending measure for the PCP, at least in the short run. If the primary care physician successfully helps the patient reduce the frequency with which they use the emergency room, then the *payer's spending for that patient* will show a decrease, but if that patient is still using emergency room services at higher rates than other patients, the overall *average spending attributed to the PCP* will increase.

  ***EXAMPLE:*** *In Figure 3, a hypothetical primary care (PCP) practice has 1,000 patients. 10% of the patients use the emergency room frequently for problems that could have been addressed by the primary care practice. Although the high ER utilizers have visited the practice in the past, they do not make any visits to the practice during the current year and they make an average of 6 visits to the emergency room during the year. The following year, the PCP practice makes an effort to see the high ER utilizers in the office and the practice succeeds in cutting the rate of ER visits by those patients in half. In the first year, none of the high utilizers would have been attributed to the practice (since the patients made no visits to the practice), but in the second year, all of the high utilizers would be attributed to the practice. Even though the total spending on the patients decreases by 41% as a result of the primary care practice's efforts, the spending attributed to the PCP practice quadruples.*

- **The patients most in need of care coordination will be excluded, and providers who provide coordination may be penalized.** The larger the number of physicians and other providers who are involved in a patient's care, the less likely those patients are to be assigned to any physician under attribution rules that try to identify a single physician who delivered a minimum proportion of the total care. Yet these patients are also likely to be among those where the opportunities to reduce spending are the highest. For example, a study done by Robert Houchens and colleagues for the Medicare Payment Advisory Commission found that in six different metropolitan areas studied, 12-22% of total spending for Medicare beneficiaries was

## FIGURE 3
### Differences in Attributed Spending if High ER Utilizers Are Seen By a Primary Care Practice

| | Year 1 | | Year 2 | | |
|---|---|---|---|---|---|
| | Low ER Utilizers | High ER Utilizers | Low ER Utilizers | High ER Utilizers | Change |
| **Number of Patients** | 900 | 100 | 900 | 100 | |
| **PCP Office Visits** | | | | | |
| Visits Per Year Per Patient | 1 | 0 | 1 | 2 | |
| Payment Per Visit | $70 | $70 | $70 | $70 | |
| Total PCP Spending | $63,000 | $0 | $63,000 | $14,000 | |
| **Emergency Room Visits** | | | | | |
| Visits Per Year Per Patient | 0 | 6 | 0 | 3 | |
| Payment Per Visit | $750 | $750 | $750 | $750 | |
| Total ER Spending | $0 | $450,000 | $0 | $225,000 | |
| **Total Spending** | $63,000 | $450,000 | $63,000 | $239,000 | |
| **Average Spending Per Patient** | **$513** | | **$302** | | **-41%** |
| Attributed Patients | 900 | 0 | 900 | 100 | |
| Attributed Spending | $63,000 | $0 | $63,000 | $239,000 | |
| **Attributed Spending Per Patient** | **$70** | | **$302** | | **+331%** |

associated with episodes in which ten or more physicians were involved, and the spending per episode in these cases averaged between $8,500 and $11,000, compared to an average of a few hundred dollars for the cases when only one physician was involved.[38]

Moreover, the minimum threshold in the attribution rule creates a perverse incentive – if one physician increases their involvement in a complex patient's case in order to provide coordination, the *total* spending for the patient might then be attributed to that physician, and the relatively high expenses associated with the patient (even if they were lower than they would have otherwise been thanks to the physician's greater involvement) would make that physician appear worse on a measure of average spending per patient.

- **Providers can be penalized for keeping patients healthy.** At the other extreme, the healthiest patients will also be less likely to be attributed to any physician because the attribution rules are based on the billable healthcare services that a patient receives, not based on which physician the patient would identify as their "regular physician." If a patient who is healthy does not need any healthcare services at all and does not have any billable visits with their physician during the measurement period, there will be no claims connecting the patient to the physician, and the attribution methodology will not assign the patient to the physician.

  Here, the perverse effect is that a physician or provider organization that is more successful at keeping patients healthy might be classified as "inefficient," because the healthy patients with no spending would not be attributed to the provider and therefore not included in a measure of average spending per attributed patient, thereby making the average spending per attributed patient for that provider look artificially high.[39] For example, in a study by J. Michael McWilliams and colleagues, the patients least likely to be assigned to the same ACO in two successive years were those with the smallest number of chronic conditions.[40]

***EXAMPLE:*** *Figure 4 shows two hypothetical primary care practices (PCPs) with 1,000 patients. The healthiest patients ("More Healthy") visit the primary care practice every other year for a visit with the physician, and otherwise deal with their healthcare needs over the phone or through email; spending on these patients outside of the PCP practice totals approximately $600 per year. A less healthy group of patients ("Less Healthy") visits the PCP annually, and spending on services delivered to them by providers outside the PCP practice totals about $1,200 per year. The least healthy group of patients ("Least Healthy") visits the*

*PCP three times per year, and spending for them totals about $5,000 per year.*

*PCP Practice #2 keeps its patients healthier, so that 50% are in the "More Healthy" category, whereas only 10% of the patients in PCP Practice #1 are in the "More Healthy" category. Moreover, PCP Practice #2 orders fewer unnecessary tests so that spending per year per patient for each group of patients is 5% lower for its patients than for similar patients in PCP Practice #1. As a result, the actual total spending per patient is 14% lower in PCP Practice #2 than in PCP Practice #1. However, the spending per attributed patient is 4% higher in PCP Practice #2. Half of the healthiest patients are not attributed to the practice because they did not have an office visit with the practice during the year. Because PCP Practice #2 has more such patients, a higher proportion of the patients attributed to Practice #2 are the least healthy patients, making the practice appear more expensive.*

Paradoxically, many programs designed to help primary care practices become "patient-centered medical homes" could result in even fewer patients being assigned to the primary care practice. Medical home programs typically provide an additional payment to a primary care practice to support services that are not currently billable through the fee-for-service system. If those services enable the patient to stay healthier and enable the primary care practice to manage the patient's care with fewer office visits (e.g., by having nurses visit them in their homes), the patient will have fewer billable office visits on claims forms, and as a result, the patient will be less likely to be attributed to the practice by attribution formulas based on billable visits. The patients who *do* come to the physician's office will then be the sickest patients who require more services, and they will be the only patients who will be assigned/attributed to the practice. As a result, measures of quality and cost for the practice that are based on *attributed* members might appear *worse*, even though measures based on *all* of the practice's patients might be *better*. Moreover, if the amount of the non-visit-based payment itself is based on the number of attributed patients, the practice faces a Catch-22: if the practice uses the non-visit-based payment to manage a patient's care without office visits, the lack of visits means the patient

## FIGURE 4

### Differences in Attributed Spending if Patients Are Healthier and Need Fewer PCP Visits

| | PCP Practice 1 | | | | PCP Practice 2 | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | More Healthy | More Healthy | Less Healthy | Least Healthy | More Healthy | More Healthy | Less Healthy | Least Healthy | Difference |
| **Number of Patients** | 50 | 50 | 500 | 400 | 250 | 250 | 100 | 400 | |
| **PCP Office Visits** | | | | | | | | | |
| Visits Per Patient in Current Year | 0 | 1 | 1 | 3 | 0 | 1 | 1 | 3 | |
| Payment Per Visit | $70 | $70 | $70 | $70 | $70 | $70 | $70 | $70 | |
| Subtotal | $0 | $3,500 | $35,000 | $84,000 | $0 | $17,500 | $7,000 | $84,000 | |
| **Other Spending** | | | | | | | | | |
| Spending Per Patient | $630 | $630 | $1,260 | $5,040 | $600 | $600 | $1,200 | $4,800 | -5% |
| Subtotal | $31,500 | $31,500 | $630,000 | $2,016,000 | $150,000 | $150,000 | $120,000 | $1,920,000 | |
| **Total All Spending** | $31,500 | $35,000 | $665,000 | $2,100,000 | $150,000 | $167,500 | $127,000 | $2,004,000 | |
| **Average Spending Per Patient** | **$2,832** | | | | **$2,449** | | | | **-14%** |
| Attributed Patients in Current Year | 0 | 50 | 500 | 400 | 0 | 250 | 100 | 400 | |
| Attributed Spending in Current Year | $0 | $35,000 | $665,000 | $2,100,000 | $0 | $167,500 | $127,000 | $2,004,000 | |
| **Attributed Spending Per Patient** | **$2,947** | | | | **$3,065** | | | | **4%** |

would no longer be attributed to the practice and the practice will lose the non-visit-based payment for that patient.

If each health plan is generating its own measures of spending, the likelihood of a patient being attributed to a physician or provider organization also depends on the structure of the local insurance market. If there are multiple health plans competing for patients, and if patients (or their employers) switch insurance plans frequently (a situation commonly referred to as "churn"), the percentage of patients who will be attributed to any provider by each health plan will be lower than in a market with fewer plans or less churn. This is because each health plan will run its attribution model using its own claims data, and a new member of the health plan will not have any visits with any provider in that health plan's claims history when the patient first joins the health plan. Both the patient and their physician may believe the physician is continuing to care for the patient independent of the change in the health insurance carrier, but the patient's new health plan won't "know" the physician is caring for the patient until the next time the patient visits the physician. The healthier the patient, the longer it will take for this attribution to occur. Some multi-payer medical home projects have reported that 30-40% of the patients in a primary care practice are not being attributed to them.[41] This problem can be reduced by generating spending measures using all-payer data, but it cannot be solved by individual health plans acting on their own. Even multi-payer data will miss patients who are uninsured for a portion of time and who pay for services out of pocket.

A similar problem also occurs in episode calculations. Because of concerns about the accuracy of diagnosis codes on physician claims (see Section III-D for a more detailed discussion of issues regarding claims-based diagnosis codes), some episode groupers require that there be two or more claims showing the same diagnosis code before the grouper will determine that an episode for that condition exists. If a patient has not had multiple physician visits within the relevant timeframe that were billable to the payer or other entity calculating the measure (or if the visits were unrelated to the condition and therefore did not record the diagnosis code relevant to the episode), then the episode will not be created and none of the costs associated with the episode may be attributed to the physicians who provided services within that episode.

## B. Providers Cannot Control All of the Services and Spending Assigned to Them

Even when a patient's spending is attributed to a physician or provider organization, that does not mean that the physician or organization could control or even influence all of the services that generated that spending.

### Inability to Control All Services Patients Receive

For example, in the Mehrotra study cited earlier, for the patients whose total spending was attributed to a single physician, between 45% and 63% of the total spending on physician services to those patients was billed by physicians *other* than the physician to whom the patient was assigned.[42] In the McWilliams study cited earlier, of the patients assigned to an ACO, 67% of those patients' office visits with specialists were provided by specialists outside of the ACO.[43]

Under typical fee-for-service payment systems and patient benefit designs, patients have the freedom to see multiple physicians and other providers, even for the same health problems; each of the physicians and other healthcare providers involved in the patient's care makes their own decisions about how to deliver care; and those individual decisions can independently increase or decrease the number or cost of services the patient receives. Assigning accountability for total spending to a single physician or provider organization may, in theory, give that physician or provider an "incentive" to take actions to reduce spending by other physicians and providers,[44] but if the physician cannot reasonably expect to influence that other spending, the accountability system may have little impact on spending while creating significant frustration for providers. If all of the physicians involved in a patient's care are part of the same physician group or health system, then the physicians have the organizational ability to control their combined spending, but that is not the case if the physicians practice independently.

> ***EXAMPLE:*** *Assume that during a healthy patient's visit to his primary care physician, the PCP recommends that the patient receive a screening colonoscopy. The colonoscopy is performed by a gastroenterologist from a separate medical group, not by the primary care physician. The gastroenterologist has choices about how and where to perform the colonoscopy that can affect how much the patient's insurance company pays for the colonoscopy. Both the payment for the patient's primary care visit and the payment for the colonoscopy will be included in the total spending for that patient and all of that spending would be attributed to the PCP. However, it is not reasonable to hold the PCP accountable for the total spending, since the PCP has no control over how much the gastroenterologist is paid for the colonoscopy or whether the gastroenterologist does the colonoscopy in a way that makes it more expensive. The decision as to which gastroenterologist is used may be up to the patient or the patient's health plan, not the PCP, and in some communities, there may be only one gastroenterologist or gastroenterology practice the patient can use. One would not want the PCP to avoid recommending an appropriate colonoscopy simply because the cost of colonoscopies in the community is higher than in other communities or because success in getting patients to have colon cancer screenings increases the PCP's spending per patient. (It would also not be reasonable to hold the gastroenterologist solely accountable for the colonoscopy spending, when it was the PCP who recommended that the patient receive the colonoscopy.)*

### Inadequate Resources or Expertise to Manage All of a Patient's Conditions and Services

Attributing all of a patient's services and spending to a primary care physician implicitly assumes that a primary care physician can and should manage and treat all of the

patient's health conditions. However, for many health conditions, such as cancer or pregnancy, a specialist will typically provide all of the treatment and management, not a primary care physician, and it would be inappropriate to expect the primary care physician to interfere with or second-guess the treatment decisions made by that specialist.

Even if a patient has health conditions that can generally be managed by a primary care practice, such as COPD, diabetes, or hypertension, a subset of patients will have multiple conditions that require much more complex and intensive care management and care coordination services. Most primary care practices are not paid adequately or at all to provide these kinds of care management and care coordination services, so it is inappropriate to hold the practices accountable for the high spending associated with the patients who could have benefited from such services, and if they are held accountable for the spending, it could discourage the primary care practices from becoming involved at all in those patients' care.

## Inability to Control Services Before Becoming Involved in a Patient's Care

It is particularly inappropriate to expect a physician or other provider to influence services a patient received before the physician first became involved in a patient's care. However, attributed spending measures are typically based on an arbitrarily defined retrospective "lookback" period (e.g., a calendar year) and the attribution methodologies typically do not distinguish the sequence of services that occurred during that year. As a result, although an individual patient would only be assigned to a physician based on services that occurred *after* the patient began seeing the physician, if a physician met the threshold to have the patient assigned, then *all* of the services the patient received, including those *prior* to their initial visit with that physician, would also be attributed to that physician. Similar to earlier examples, this can create a perverse incentive for a physician not to become involved with a patient who already incurred significant healthcare costs earlier in the year, even though these are the patients who may most need additional help.

*EXAMPLE: A patient has not visited any physicians or received any healthcare services for several years. At the beginning of the calendar year, the patient has a serious heart attack and receives cardiac bypass surgery. When the patient is discharged, he is encouraged to begin seeing a primary care physician. The patient chooses a primary care practice and begins visiting the practice both for care of his heart disease and for more general preventive care. At the end of the year, the patient's health plan assigns the patient to the primary care practice he has been using because the practice delivered all of the primary care services to the patient during the year. The cost of the hospitalization and surgery is then attributed to the primary care practice, even though it occurred before the patient's first contact was made with the primary care practice, and even though the practice had no influence over the services delivered during that hospitalization.*

Because spending on all services of any kind are typically included in accountability systems, even spending that occurs *after* a physician's initial involvement with a patient can be inappropriately assigned to the physician. For example, if a patient is seeing their primary care physician regularly, but suffers serious trauma in an automobile accident, the cost of the trauma care would be assigned to the PCP, even though there was nothing the primary care physician could have done to prevent the accident, reduce its severity, or control the costs of the trauma care the patient received. A patient may also see a primary care physician solely for minor acute needs, and the physician may not be given the time or a reason to evaluate the patient for the presence of other conditions. If the patient is later hospitalized for a problem related to one of those conditions, the cost of the hospitalization would be attributed to the primary care physician even though he or she had never been involved in that aspect of the patient's care.

## Inability to Control All Services Within Episodes

Assigning accountability for spending based on episodes mitigates but does not eliminate this problem. In many cases, a patient will receive services from multiple physicians during a single hospitalization or longer episode of care. For example, in the Mehrotra study cited earlier, between 27% and 51% of the total spending on physician services *within episodes* was delivered by physicians *other than the physician to whom the episode was assigned*.[45]

In most cases, none of the individual physicians involved in a patient's care has any authority over the services delivered by the others, even within a specific episode. Moreover, the longer the time period associated with an episode of care, the more providers will be involved and the less influence any of them will have over the others, and some may not even be aware of the others' involvement.[46] For example, if a hip surgery patient chooses to go to a rehabilitation facility after discharge from the hospital, develops an infection because of poor wound care in the rehabilitation facility, and has to be readmitted to the hospital to treat the infection, the costs of the hip surgery, the rehabilitation services, and the treatment for the infection would typically be included in the hip surgery episode,[47] but it would be difficult to argue that the surgeon or any other provider should be held solely accountable for all of those costs.[48]

Consequently, although dividing total spending into episodes can make attribution seem more rational than simply using total spending, the attribution methodologies still result in providers being held accountable for aspects of spending within the episodes that they cannot control.

## Lack of Actionable Information on Services

The typical reports on spending that are given to physicians and providers do not enable the providers to determine which aspects of the spending resulted from the services they delivered or could have influenced. Figure 5 shows an example of a report CMS provided physicians through its Quality and Resource Utilization Reports (QRUR) program that is similar to what many commercial

## FIGURE 5

### 2010 Per Patient Costs in 4 States (Iowa, Kansas, Missouri, Nebraska) For Medicare Patients Whose Care was "Directed" by a Cardiologist

(SOURCE: 2010 CMS QRUR Reports)

| Type of Service Received by Patient | Average Risk-Adjusted Spending Per Patient | Likely Questions by the Cardiologists Receiving the Report |
|---|---|---|
| **Services Provided by the Cardiologist** | | |
| E&M Services (Office Visits) | $222 | Did these visits occur before or after the other spending? |
| Procedures | $102 | Where were these procedures performed? |
| Subtotal | $324 | |
| **Services Provided by Other Physicians** | | What kinds of physicians are these? |
| E&M Services (Office Visits) | $466 | Were these visits based on referrals by the cardiologists? |
| Procedures | $414 | Were these procedures ordered by the cardiologists? |
| **Hospital Services** | | |
| Inpatient Hospital Stays | $3,891 | Were these hospital stays for heart problems? |
| Clinic or Emergency Visits | $364 | Were these visits for heart problems? |
| Hospital Outpatient Procedures | $1,431 | Were these procedures for heart problems? Who performed them? |
| Hospital Outpatient Laboratory and Testing | $376 | Were these tests for heart problems? Who ordered them? |
| Hospital Outpatient Imaging | $436 | Were these imaging studies for heart problems? Who ordered them? |
| **Services in Ambulatory Settings** | | |
| Laboratory and Other Tests | $240 | Were these tests for heart problems? Who ordered them? |
| Imaging Services | $364 | Were these imaging studies for heart problems? Who ordered them? |
| Durable Medical Equipment | $317 | Was this equipment related to cardiac care? |
| **Post-Acute Care Services** | | |
| Skilled Nursing Facilities | $762 | Did these services follow hospital stays for heart problems? |
| Psychiatric, Rehab or Other LTC | $162 | Did these services follow hospital stays for heart problems? |
| Home Health | $281 | Did these services follow hospital stays for heart problems? |
| **Other Services** | | |
| All Other Services | $840 | Were these services related to heart problems? |
| **Total Risk-Adjusted Per Capita Spending** | **$10,667** | |

payers provide to physicians.[49] The example shows the average total spending on care for patients in four states for whom a cardiologist was determined to have "directed" the patient's care, meaning (according to the rule CMS used) that the cardiologist billed 35% or more of the outpatient evaluation and management visits with the patient during the course of the year. The spending shown on the report includes all services of any type received by the patient during the year, not just cardiology-related services. For example, if the patient was in an auto accident and suffered broken bones, the costs of repairing the fractures would be included in the total spending in the report even though the cardiologist may have had no involvement with the patient's care for those injuries. If the patient was being treated for colon cancer, the costs of the surgery, radiation, and/or chemotherapy for the cancer would be included in the total spending in the report given to the cardiologist, even though the cardiologist would not have been responsible for any of these services.

Although a breakdown of the data into service categories is provided in Figure 5, the categories used do not answer the kinds of questions the cardiologist would likely have regarding which services the cardiologist was involved in or what services were cardiology-related. Even though the cardiologist billed 35% or more of the *evaluation and management* services for the patients included in the report, the total services that were directly billed by the cardiologist for those patients represent a mere 3% of the *total spending* on the patients ($324 of the $10,667 total per patient spending). Some of the cardiologist's services likely resulted in a portion of the remaining 97% of the spending that is reported in the other categories (e.g., the payments to hospitals or other facilities for tests ordered or procedures performed by the cardiologist), but services from other physicians likely did as well. In fact, the report shows that other physicians directly billed for 2.7 times as much as the cardiologist did (other physicians billed an average of $466 for office visits and $414 for procedures, for a total of $880 vs. the $324 billed by the cardiologist), but the report does not indicate how much of the spending on hospitals and other non-physician providers resulted from what those other physicians may have done or ordered, or even what kinds of physicians or services they were.

Whether one is trying to merely inform providers about ways to reduce spending or hold them accountable for spending, if information on spending is not tabulated in a way that distinguishes spending that providers can and cannot influence and that helps providers understand what contributions they made to total spending and what actions they can take to change, the information will be of little or no value, and any accountability program based on this information could be ineffective at best and could potentially be counterproductive for patients and payers.[50]

## C. Providers Are Not Attributed Many Services They Do Provide

### Services Measured But Not Attributed to the Responsible Provider

The obverse of the problem described in the previous section – attributing a service delivered by one provider to a different provider who could not control it – is that the service is not attributed to the provider who could control it.

In the Pham study cited earlier, the patients who were assigned to primary care physicians only represented 39% of the Medicare patients that those physicians actually saw during the course of the year, i.e., the majority of a primary care physician's patients were not attributed to them.[51] In the Mehrotra study, no more than 23% of the physicians' own billings were attributed to them when patients were assigned to physicians based on total spending and no more than 60% of physicians' own billings were attributed to them under any of the attribution rules studied, including when episodes were being attributed instead of total spending.[52] In the study by McWilliams and colleagues, only 38% of Medicare spending on outpatient services billed by an ACO was associated with the patients assigned to it under Medicare attribution rules.[53]

> *Most of the spending attributed to physicians results from services delivered by other physicians, and most of the spending on services physicians deliver is not attributed to them.*

This means that the spending analysis for a provider will not be based on the full range of patients they care for or the full range of services they provide, but rather on a non-random subset of those patients. The subset of patients and services that are assigned to them will depend on the number and types of the other providers and services involved in their patients' care, which in turn will depend on the community where they practice, the types of patients they see, the types of insurance those patients have, etc. It might be possible to make adjustments to the spending measures to reduce or eliminate the biases this causes, but this opportunity does not exist when the attributed spending measures are simply calculated and used to adjust payments formulaicly.

Of particular concern is that patients who experience serious complications from treatment may be less likely to be attributed to the physician or other provider whose treatment caused the complication. If a complication is serious, such as a severe infection, the treatment *for* the complication may require many more services than were involved in delivering the treatment that *caused* the complication, and the physicians and other providers who treat the complication may be different than those who delivered the initial treatment. Since most attribution rules for episodes assign responsibility to the physician who delivered the largest number of services or most ex-

pensive services in the episode, an episode in which a serious complication occurred could be assigned to the physician who *treated* the complication, not the physician who *caused* the complication. This would result in attributing a misleadingly low amount of spending to the physician who caused the complication and a misleadingly high amount of spending to the physician who treated the complication.

*EXAMPLE: Figure 6 shows two hypothetical patients receiving bowel surgery in a hospital. The surgery for Patient 1 is successful. All of the physician fees during the hospital stay are those billed by the surgeon, and so the surgeon is attributed the spending for Patient 1. The surgery for Patient 2 is less successful – the patient develops a severe infection following surgery which requires the patient to stay in the hospital an extra week. A hospitalist successfully treats the infection and the patient is able to be discharged. During the extra week the patient is in the hospital, the hospitalist's fees cumulate to more than the surgeon's fee for the surgery. Since the hospitalist is responsible for the majority of the physician fees during the stay, the hospitalist is attributed the spending for Patient 2.*

In addition, there have been widely-publicized cases in which physicians and other providers were delivering large numbers of services inappropriately or fraudulently. In order to hold these providers accountable for this, they first need to be identified, and analyses of the services that physicians have delivered provide one way to do this.[54] In most cases, these physicians are not the patient's primary care physician and it is unlikely that they will be providing the majority or plurality of the services or spending for any individual patient. As a result, the attribution methodologies described in Section II would assign their spending to a different physician, such as the patient's primary care physician. Since the high spending by these physicians would be hidden in the spending totals assigned to a large number of PCPs, it would be difficult or impossible to identify and reduce the unnecessary spending they are causing, and some PCPs may be inappropriately penalized if they happen to have a large number of patients who choose to obtain services from another provider who delivers services inappropriately.

## Services Not Measured At All

A separate problem is that many spending analyses do not include the spending for all types of services that patients receive and this also introduces biases into the measurement of total spending for providers.

For example, as noted in Section II, spending measures that are described as "total spending" or "total cost" frequently exclude spending on prescription medications, despite the fact that spending on drugs represents over 10% of healthcare spending in the U.S.[55] Because health insurance for prescription medications is frequently administered separately from insurance for medical services, and because some patients have medical insurance but not prescription coverage, spending tabulations using claims data may exclude information on medication spending. For example, the "Total Per Capita Cost" measure used by CMS in its Physician Value-Based Payment Modifier does not include Part D (drug) spending, only Part A (inpatient services) and Part B (outpatient services).

Since prescription medications play a key role in keeping patients healthy and in enabling them to recover from various health problems, higher spending on drugs may reduce spending on services such as inpatient care, and vice versa. As a result, spending totals for physicians who prescribe more drugs but use fewer other services will appear artificially low compared to other physicians. In addition, some types of drugs are paid for through the patient's medical insurance while other types of drugs are paid through prescription insurance, so spending tabulations for physicians whose patients differ in the types of drugs they need will not be comparable if the spending under prescription insurance is not included.

### FIGURE 6
### Attribution of Episodes With and Without Complications for Two Hypothetical Patients

| | Patient 1: Successful surgery | | Patient 2: Surgery with post-op infection | | |
|---|---|---|---|---|---|
| | SURGEON<br>Surgery for Bowel Obstruction | Total | SURGEON<br>Surgery for Bowel Obstruction | HOSPITALIST<br>Treatment of Infection from Surgery | Total |
| **Physician Fees** | $750 | | $750 | $834 | |
| **Hospital Payment** | $6,500 | | $6,500 | | |
| **Total Spending** | $7,250 | $7,250 | $7,250 | $834 | $8,084 |
| **Share of Physician Services** | **100%** | | **47%** | **53%** | |
| **Attribution:** | **X** | | | **X** | |

*EXAMPLE: If patients with cancer are treated using infused chemotherapy, those drugs will typically be paid for through medical insurance and the costs of those drugs will be included in claims-based spending tabulations, but if the patients are treated with oral chemotherapy, those drugs will typically be paid for through pharmaceutical insurance and the costs of those drugs will be excluded from claims-based spending tabulations. As a result, an oncologist who has more patients that can be treated with oral chemotherapy will appear to be "lower cost" than an oncologist whose patients can only be treated with infused medications.*

## D. Spending Measures Do Not Distinguish Appropriateness of Services

Even if one identifies services that providers can control or influence, it is important to ensure that efforts to encourage lower spending do not harm patients. It is widely agreed that there are significant opportunities to reduce healthcare spending without harming patients because many of the services that patients receive today are unnecessary, and some are even harmful.[56] However, typical accountability systems combine spending on recommended services, inappropriate services, overused services, avoidable services, and typical services into a single total. As a result, a provider who does a better job of delivering recommended services could be measured as having higher spending than a provider who fails to deliver recommended services or a provider who delivers services that are less expensive but inappropriate for the patient. This could have the unintentional side effect of encouraging providers to stint on desirable care to patients in order to reduce the total amount of spending.[57]

*EXAMPLE: Figure 7 shows the total cost per patient for patients between 50 and 75 years of age in three hypothetical physician groups that include both primary care physicians and gastroenterologists. In each group, the primary care physicians see each patient in the office an average of twice per year at a cost of $70 per visit. Preventive care guidelines recommend that patients in this age range receive a screening colonoscopy every 10 years.*

- *In Physician Group #1, none of the patients receive the colonoscopy that is recommended by guidelines.*
- *In Physician Group #2, only one-fifth of the patients who should have a colonoscopy each year get one. The gastroenterologists in Physician Group #2 perform all of their colonoscopies at a hospital at a cost of $820 (a $220 payment to the gastroenterologist and a $600 payment to the hospital).*
- *In Physician Group #3, all patients receive colonoscopies in accordance with the guidelines.[58] The gastroenterologists in Physician Group #3 perform their colonoscopies in the office rather than at a hospital, at a total cost of only $400 per colonoscopy, less than half as much as the cost of a colonoscopy performed by Physician Group #2.*

*The average spending per patient for primary care and gastroenterology services by Group #1 is $140, the average spending for Group #2 is $156.40, and the average spending per patient by Group #3 is $180, meaning that Group #3, which is the most successful in using evidence-based care and providing colonoscopies at the lowest cost per colonoscopy, actually has the highest spending per patient. In fact, Group #3 is more than one standard deviation above the average spending for the three practices. As a result, it would be labeled as a "high cost" practice in many cost tiering systems and its payments would be reduced under the Medicare Value Based Payment Modifier.[59]*

## E. Risk Adjustment Systems May Not Adequately Adjust for Patient Needs

All else being equal, a physician or hospital that treats sicker patients will likely spend more per patient because the patients will need more services. It would be inappropriate to say that a provider is "more expensive" than another if that provider's patients have more health problems, more severe health problems, or other relevant differences from the patients cared for by other providers. As explained in Section II, most accountability systems use some type of risk adjustment system in an effort to separate differences in spending due to differences in patient needs from differences in the way providers deliver care. Unfortunately, the risk adjustment systems used in most accountability systems may not effectively separate differences in patient needs from differences in the way providers deliver care, for several reasons:

- Most risk adjustment systems are designed to predict *spending* on patient care, not adjust for differences in patient *needs*.
- Most risk adjustment systems use *historical* information on patient characteristics, not the most current information on health problems that affect the services patients need.
- Most risk adjustment systems produce a *single* risk score for a patient, even though different aspects of patients' health conditions will have different implications for the care delivered by providers in different specialties.
- Most risk adjustment systems use information available in *claims data* that does not completely or accurately measure differences in patient health needs.
- Most risk adjustment systems give little or no consideration to factors *other than health status* that can affect patient needs.

For example, the Hierarchical Condition Category (HCC) system used by CMS for risk adjustment is designed to predict future spending, not measure current patient needs.[60] It changes the weights for individual conditions significantly from year to year based on which factors achieve the best results in regression-based predictions of actual spending in the most recent year, not based on changes in clinical evidence about what patients need; it explicitly gives zero weight to many acute conditions, even though these conditions would likely result in a need for

## FIGURE 7
### Comparison of Spending Per Patient For Three Hypothetical Physician Groups

| | Physician Group 1 | Physician Group 2 | Physician Group 3 |
|---|---|---|---|
| **Number of Patients Ages 50-75** | 2,000 | 2,000 | 2,000 |
| Primary Care Visits Per Patient Per Year | 2 | 2 | 2 |
| Total Number of Primary Care Visits | 4,000 | 4,000 | 4,000 |
| Payment Per Primary Care Visit | $70 | $70 | $70 |
| **Total Payments to Primary Care Physicians** | **$280,000** | **$280,000** | **$280,000** |
| Proportion of Patients Receiving Recommended Colonoscopies | 0% | 20% | 100% |
| Total Number of Colonoscopies Performed (1 Every 10 Years) | 0 | 40 | 200 |
| Proportion of Colonoscopies Performed in Office | 0% | 0% | 100% |
| Payment to Gastroenterologist for Office Colonoscopy | | | $400 |
| **Total Payments to Gastroenterologists for Office Colonoscopies** | **$0** | **$0** | **$80,000** |
| % of Colonoscopies Performed at Hospital Outpatient Center | 0% | 100% | 0% |
| Payment to Gastroenterologist for Outpatient Hospital Colonoscopies | | $220 | |
| Payment to Hospital for Outpatient Colonoscopy | | $600 | |
| **Total Payments to Gastroenterologists for Hospital Colonoscopies** | **$0** | **$8,800** | **$0** |
| **Total Payments to Hospital for Hospital Colonoscopies** | **$0** | **$24,000** | **$0** |
| Total Spending | $280,000 | $312,800 | $360,000 |
| **Average Spending Per Patient** | **$140.00** | **$156.40** | **$180.00** |
| **Overall Average** | | **$158.80** | |
| Standard Deviation | | $20.11 | |
| **Average +/- 1 Standard Deviation** | **$138.69** | | **$178.91** |
| ***Physician Group Rating on Spending Per Patient*** | ***Average*** | ***Average*** | ***High*** |

*"High" = Per Patient Spending > Average + 1 Standard Deviation*

*"Low" = Per Patient Spending < Average + 1 Standard Deviation*

services during the year in which they occurred and could also affect service needs in future years; it uses only diagnosis information from claims data; and it does not consider many factors other than health conditions that can affect patient needs.

## Risk Adjustment Systems Are Designed to Predict Spending, Not to Measure Differences in Patient Needs

Most risk adjustment systems were not developed specifically for use in measurement and accountability systems, but instead were developed to help health plans set premiums for groups of patients or to enable purchasers and payers predict how much will be spent on healthcare services for a particular patient population. A risk adjustment system will assign a higher risk score to a patient if the amount that is *typically spent* on similar patients is higher, even if those patients did not actually *need* all of the services they received. Consequently, using risk scores calculated as they are today can actually reinforce inappropriate spending, penalize efforts to reduce underuse, and cause providers to focus spending reduction efforts on the wrong patients.

***EXAMPLE:*** *Consider two hypothetical groups of patients with two different sets of health problems. Assume that the best available evidence suggests that the "right" care for each group of patients would require roughly the same amount of spending (i.e., the specific services would differ because the health problems differ, but the total cost of each set of services would be similar). Assume further that (1) most providers deliver much more care than evidence indicates is needed to patients with the first health problem (i.e., there is "overuse") and/or those patients experience many avoidable complications, so that the actual spending for those patients is higher than necessary,*

*and (2) most providers deliver less than the recommended level of care to patients with the second health problem (i.e., there is "underuse"), so the spending for the second group of patients is lower than would be considered ideal. A typical risk adjustment system would classify the first set of patients as "higher risk" than the second set of patients, because the risk score for the patients is based on how much is actually spent on treating them, not on how much care the patients truly need. Similarly, the patients in the second group would be classified as "lower risk" simply because there has been less spending on them in the past than they needed. If spending is adjusted using risk scores calculated in this way, the unfortunate result will be that (1) any providers who actually deliver the recommended level of care to the second group would appear to be "expensive" or "inefficient," (2) providers in the first group will be ranked as "efficient" even if they use unnecessary services, and (3) providers could be encouraged to focus spending reduction efforts on the second group rather than the first group, even though there is overuse in the first group and underuse in the second group.*

If overuse of services were directly correlated with appropriate use (e.g., if the unnecessary proportion of the total services that patients received was the same for both patients with low levels of need and high levels of need), then the factors and formula that a risk adjustment system used to predict *total* spending might also be a reasonable way of predicting *appropriate* spending. However, if overuse occurs independently of appropriate use, as studies seem to suggest, then risk adjustment systems that are better at predicting differences in *total* spending across patient populations may actually be *worse* at predicting differences in actual patient needs. As a result, risk-adjusted spending measures can provide misleading information about where opportunities to reduce spending exist.

*Current risk adjustment methodologies can reinforce inappropriate spending, penalize efforts to reduce underuse, and cause providers to focus spending reduction efforts on the wrong patients.*

Most risk adjustment systems are developed using statistical techniques such as linear regression analysis that do not select which patient characteristics to use and how to weight those characteristics based on clinical judgments about which patients would appropriately require more services and more spending, but rather based on what characteristics are statistically most successful in predicting spending in the year(s) of data on which the regression analysis was performed. These regression-based risk adjustment systems are very problematic for use in comparing spending over time. Because the weights used in the risk adjustment score are determined statistically, not based on scientific evidence or appropriate use guidelines as to what kind of care is appropriate, the weights can change significantly from year to year based simply on what combination of weights and variables best predicted spending for the most recent year(s) of data used to generate regression-based risk scores. This means that patients can have different risk scores from year to year even if the patient's health conditions did not change and even if there had been no change in medical evidence to suggest that more or less care was appropriate for patients with those health conditions. (For example, CMS issues a new set of weights for its HCC risk adjustment methodology every year, and there are often significant changes from one year to the next.)

In addition, unless the weights are determined nationally across all payers, different weights might be calculated by different payers for similar groups of patients, and so a provider who is delivering the same care to all of their patients that have the same health problems might receive different risk-adjusted spending scores from different payers.

## Most Risk Adjustment Systems Use Historical Information on Patient Characteristics, Not Current Information

Since risk adjustment systems are typically developed to predict *future* spending, not to make adjustments to *current* levels of spending for accountability purposes, the risk scores are calculated based only on health problems the patient had in *previous* years in order to predict spending in the current year or future years.

However, since the risk scores generated by these "prospective" risk adjustment systems ignore health problems the patient developed in the current year, they will underestimate the amount of spending that the patient would need in the current year.[61] In fact, prospective risk adjustment systems are explicitly designed to ignore temporary conditions, such as short-term acute conditions, that would cause spending to be higher in the year in which they occur but would be unlikely to cause higher spending in future years. For example, the Medicare HCC risk adjustment system explicitly gives zero weight to many acute conditions, even though these conditions would likely result in a need for services during the year in which they occurred and could also affect service needs in future years.

As a result, two providers may each have patients with similar average *prospective* risk scores, but if one provider's patients developed more new health problems in the current year than the other's did, the first provider would appropriately have more spending than the second provider in the current year. However, since the risk scores for the patients would not change until the following year, if at all, the first provider would appear to have higher risk-adjusted spending than the second provider in the current year. For example, if a patient is diagnosed with metastatic cancer *this* year, a risk score calculated based on their health status *last* year will likely underestimate the significant healthcare services and associated spending the patient will need to receive this year in treating the cancer.

## A Single Risk Score Doesn't Reflect Differences in Care Needed from Different Providers

Regardless of whether the risk scores are prospective or

concurrent and even if they are designed to predict what services patients need rather than what is typically provided, a single risk score for a patient cannot accurately predict how the level of services and spending will differ for *different types of providers* who deliver different subsets of the care for the patient. Under most risk adjustment systems, two patients with very different health problems could have the same risk score if the *total* predicted spending needed to treat those different health problems was similar, but that does not mean that the spending by different physicians in different specialties should be similar.

> ***EXAMPLE:*** *In the 2014 version of the Medicare HCC risk adjustment system, a patient with colon cancer would have the same risk score as a patient who had a stroke, but one would not expect the services provided by neurologists, cardiologists, and physiatrists for the patient with a stroke to be the same as what they would provide (if anything) for a patient with colon cancer but no stroke. Similarly, an oncologist would be expected to provide oncology services to the patient with cancer but not to the patient with a stroke, even though each would have the same risk score.*

In accountability systems that assign spending to primary care physicians, the fact that the patients in two primary care practices have similar risk scores does not mean that the primary care practices have equal ability to manage the costs of those patients. In one practice, the risk score of the patients may be due to common chronic conditions such as asthma, COPD, diabetes, or heart failure of low or moderate severity that the primary care practice has the knowledge and resources to manage effectively; in the other practice, the patients' risk score may be due to the presence of conditions such as cancer, inflammatory bowel disease, macular degeneration, or mental illness that require expensive treatment and management primarily or exclusively by specialists. Although both types of patients might be attributed to the primary care practice because they visit the PCP for standard preventive care, it would be inappropriate to classify the second practice as higher spending because its patients need more specialty care.

> ***EXAMPLE:*** *In the 2014 version of the Medicare HCC risk adjustment system, patients would have similar risk scores regardless of whether they had bladder cancer, cerebral hemorrhage, COPD, diabetes with complications, heart failure, intestinal obstruction, inflammatory bowel disease, major bipolar disorder, or macular degeneration, but some of these conditions would normally be treated principally by a primary care practice and others would be treated principally by a specialist.*

Similarly, an individual patient's overall risk score would typically be calculated based on all of the health problems the patient has, but different subsets of those health problems would affect the treatment decisions of different physicians differently.

> ***EXAMPLE:*** *If a patient has both coronary artery disease and severe arthritis of the knee, both of those health conditions may need to be treated and both treatments may involve the use of expensive health services. The fact that the patient has arthritis of the knee will likely have relatively little impact on how the cardiologist treats the patient's heart disease, but the severity of the patient's heart disease may be an important consideration for a surgeon considering how to treat the patient's knee.*

Even for a primary care physician or a multi-specialty group of physicians that is trying to manage the overall care for a patient population, assigning a single risk score to patients with very different types of conditions does not provide very helpful information in determining where overuse or underuse of services may be occurring.

## Claims-Based Risk Adjustment May Not Accurately Measure Patient Health Status

Most risk adjustment systems use the information in health plan claims systems to determine how to compute a risk score for that patient. The only information on a patient's health status in claims data, however, comes from the diagnosis codes recorded on bills for services, and using this information is problematic for several reasons:

- Diagnosis codes do not always fully distinguish differences in patient conditions that can significantly influence the nature of the services they should appropriately receive. For example, in addition to the *type* of cancer a patient has (e.g., breast, colon, lung, etc.), the *stage* of cancer (e.g., whether it has metastasized to other parts of the body) has a significant impact on how it is treated by oncologists. However, neither the ICD-9 nor ICD-10 diagnostic coding systems has a method for recording the stage of cancer, only the type of cancer.
- Claims forms only allow a limited number of diagnosis codes to be recorded for any one patient. This means that for patients with multiple health problems, the full range of their needs may not be recorded.[62]
- If a patient has a chronic health problem that was diagnosed several years earlier and if that health condition is now being well-managed with medications so that the patient does not need to see a physician in person for problems related to that condition, there may not be any current claims that explicitly show that diagnosis, even though the patient still has the health condition.
- The diagnosis code may not be as precise as it should be. There are many different codes for the same general health condition that specify more detailed characteristics about the condition in individual patients; these differences in the detailed diagnosis may have a significant influence on services. However, for physician service claims, *payment* is based on the *procedure* performed, not the specific *diagnosis* applicable to a particular patient. This means that there is little incentive today for physicians to ensure that the diagnosis codes are complete or as precise as needed for accurate spending analyses, particularly when the physician is given very limited time to spend with patients. (In contrast, in the DRG system used by Medicare and many payers for hospital payment, the diag-

nosis code is a major factor used to determine the payment amount; this means that the diagnosis codes are much more likely to be complete and accurate for claims filed by hospitals.)

- In some cases, a diagnosis code is recorded not to reflect a health problem that a patient actually has, but a possible condition that a physician visit or test was attempting to verify or rule out. This means that patients who are more likely to report symptoms or have symptoms associated with multiple types of conditions may appear to be "sicker" on claims forms than other patients.

Most of these problems lead to "undercoding," rather than "overcoding," so it seems likely that in some cases, what appears to be higher-than-average risk-adjusted spending for a provider may actually be caused by having sicker patients who are not accurately classified in the risk-adjustment system. Moreover, because similar patients may be coded differently in different communities, by different providers in the same community, or even by the same provider in different years, differences in spending or changes in risk-adjusted spending may be due to differences or changes in coding, not to actual differences or changes in patient health status.

### Factors Other Than Health Status Affect Patient Needs

Differences in the number and types of diseases are not the only thing that can cause patients to have different needs for services or to experience different levels of complications with the same levels of health services. For example:

- **Differences in functional limitations and caregiver supports.** Patients who are unable to walk or drive or are unable to carry out activities of daily living will have greater difficulty caring for themselves and greater difficulty obtaining traditional office-based ambulatory care services, which can lead to increased use of more expensive healthcare services. For example, an analysis by Harriet Komisar and Judy Feder at Georgetown University found that there were hospital admissions for 34% of Medicare beneficiaries with functional limitations as well as chronic diseases, but there were admissions for only 20% of Medicare beneficiaries with 3 or more chronic conditions but no functional limitations. They also found that the majority of the beneficiaries on whom Medicare spent the most had both chronic conditions and functional limitations.[63] However, since information about functional limitations is not captured by standard diagnosis coding in claims data, it is not incorporated into most risk adjustment models. A study by Katia Noyes and colleagues found that the Medicare HCC risk adjustment model significantly under-predicted actual spending on the subset of patients with functional disabilities.[64] Patients with functional limitations who have caregivers to assist them will likely do better than patients who do not, but the availability of caregivers or other support systems is also not assessed in risk adjustment systems.
- **Differences in patient engagement.** Studies have shown that "activated patients," i.e., those with the willingness and ability to take independent actions to manage their health and care, are less likely to experience complications and less likely to use higher-cost services such as emergency rooms and hospitals. A study by Judith Hibbard and colleagues found that in one health system, spending on patients with the lowest "activation levels" was 8% higher than for patients with the highest activation levels, and for patients with some types of chronic disease, the difference was as much as 21%.[65] However, patient activation is not a factor considered in risk adjustment systems.
- **Differences in health insurance benefits.** Even if a patient has health insurance, high deductibles or high cost-sharing levels may discourage them from seeking preventive care or taking prescribed medications, leading to avoidable complications and higher overall expenses that are outside the control of their physicians and other healthcare providers. However, differences in patients' cost-sharing levels are not considered in risk adjustment systems.
- **Differences in services needed in rural areas.** Patients in rural areas will have greater difficulty using healthcare services that are delivered in traditional ways. For example, longer driving times, lack of public transportation, etc. will make it more difficult for patients in rural areas to make office visits to physicians, yet alternative means of contact – phone calls, emails, video calls – are typically not reimbursable by Medicare or commercial insurers. As a result, patients may end up using emergency rooms and other potentially avoidable services at higher rates because alternatives aren't financially feasible. However, the population density of the patient's community is not considered as a factor in risk adjustment systems.
- **Differences in socioeconomic status, education, and other factors.** Studies have shown that avoidable complications and other poor outcomes occur more frequently in patients of lower socioeconomic status, patients with lower education levels, etc., even after controlling for differences in health problems. Although some of these differences may result from the fact that these patients also have more functional limitations, lower levels of activation, different health insurance benefits, and differences in access to care, other differences may be due to factors that are not easily measured or understood.

Failing to adjust for these factors could unfairly penalize providers who care for disproportionate numbers of patients with these characteristics, which in turn could make it more difficult for those patients to find providers able or willing to care for them.

## F. Spending Measures Do Not Adjust for Structural Differences in Providers and Communities

Clearly, differences in the characteristics of patients can result in appropriate differences in the number and types

of services and the associated spending across different providers. However, differences in the characteristics of the providers, their payers, and the nature of the communities in which they are located can also result in differences in the costs and prices of services, and even in the types of services provided, that are beyond the control of providers. For example:

- **Fewer economies of scale in rural areas and small communities.** It is inherently more expensive to operate a hospital in a sparsely-populated area than in a densely-populated area. Hospitals need to have emergency services available on a round-the-clock basis and they need to have surgery suites, testing facilities, etc. available when patients need them, but a rural hospital will have fewer patients needing those services, and so the per-patient cost of services in the rural hospital will be higher than for an otherwise identical hospital located in an urban area, even after adjusting for any differences in cost-of-living. Similarly, small physician practices in rural areas will not be able to achieve the same economies of scale as practices in areas that can support multiple physicians, so if a community does not have enough patients to support more than one physician in a particular specialty, the cost of services in that specialty will inherently be higher. As a result, analyses of spending that compare rural and non-rural providers could unfairly disadvantage the rural providers.
- **Differences in the ability of patients to pay for services.** All else being equal, a hospital or physician practice serving a large proportion of patients who are uninsured or whose insurance pays low amounts for services will have to charge more to the patients who do have insurance than a provider that does not have as many uninsured patients. In any individual service line, this is simply a variant of the previous point – having fewer paying patients to cover the total costs of the facilities, equipment, and staff associated with services means that the price per patient has to be higher to cover overall costs. However, it may also be necessary for hospitals and other multi-specialty providers to charge more for the service lines that have mostly insured patients in order to cover the losses on services lines with mostly uninsured or low-paying insurers. For example, many hospitals have to make sufficient margins on services such as cardiac, oncology, and orthopedic care in order to cover losses on services such as maternity, psychiatric, and trauma care that are disproportionately used by patients with no insurance or low-paying health plans. Analyses of spending that compare providers with different mixes of insurance can penalize those providers with more uninsured and underinsured patients.
- **Differences in benefit designs and/or payment systems.** As noted earlier, differences in the type of insurance coverage that patients have can affect their ability and willingness to access different kinds of services and cause differences in spending and outcomes that are beyond a healthcare provider's control. Because different communities have different payers with different rules regarding what services are covered, there may be systematic differences in spending between providers in different communities due to differences in benefits rather than differences in the performance of the providers.

  In addition to differences in patient benefits, however, different payers also pay providers in different ways. Providers that are being paid under payment systems that give them greater flexibility to redesign care will have greater ability to reduce spending and improve outcomes than providers being paid under traditional fee-for-service systems.
- **Higher costs of teaching and research.** In addition to the costs of patient care incurred by all hospitals, teaching hospitals incur additional costs associated with training of medical students and conducting research; they also experience inefficiencies in the delivery of care because of the time and additional steps involved in training new physicians. In most cases, the principal way that teaching hospitals cover these costs and inefficiencies is through higher payments for the individual services they deliver, so their services will appear "more expensive" because of that.

## G. Significance of the Problems With Current Methodologies

It is difficult to quantitatively determine the combined impact of the problems described above on the ability to draw meaningful conclusions from specific applications of spending measures. Typically, providers are rated and compared on spending using one attribution methodology and one risk adjustment methodology, with no information about the unattributed patients, no sensitivity analysis to indicate how much the results depended on the specifications of the methodology, and little if any breakdown of the underlying data to show whether differences in spending per patient were associated with services delivered by the provider being measured versus other providers, whether there were significant differences in patient characteristics that were not distinguished by the risk adjustment methodology, etc.

### Cross-Sectional Reliability of Spending Measures

The adequacy of a spending measure for rating and rewarding providers is typically evaluated by calculating its "reliability" using a *cross-sectional* analysis that compares the variance in the average spending per patient *between providers* (the "between-provider variance") to the variance in spending *among each provider's patients* (the "within-provider variance"). If the between-provider variance is much higher than the within-provider variance, it is more likely that there are significant differences in spending between the providers during the measurement period, whereas if the within-provider variance is higher than the between-provider variance, any measured difference in the average spending between providers could merely be due to chance based on the particular set of patients the providers happened to see during the measurement period.[66]

There has been relatively little research on the reliability of spending measures, and the research that has been done

has reported conflicting results. However, common findings are that the reliability of spending measures is lower for individual physicians than for large groups of physicians, reliability is lower for measures based on small groups of patients than on large groups of patients, and reliability is lower for measures of spending in some physician specialties than others.[67] In addition, a study by Hao Yu and colleagues found that measures of service utilization for primary care physicians were more reliable for certain types of services than other services; not surprisingly, the higher reliability was associated with services typically directly delivered or ordered by primary care physicians versus services such as hospital admissions, readmissions, etc. that can be caused by a variety of factors other than the PCP's services. This implies that spending measures attributed to PCPs may be less reliable when applied to patient populations that need or receive large numbers of services outside of the PCP's control.[68]

## Temporal Reliability of Spending Measures

However, even if a cross-sectional analysis indicates that a spending measure is a *reliable discriminator* of providers' *past* performance, that does not mean the spending measure is a *reliable predictor* of their *future* performance. The fact that two providers had statistically different levels of spending *last* year does not mean that they will have different levels of spending *next* year or that their relative rankings on spending will be the same in the future, even if their underlying approach to patient care remains the same. In fact, the more that the problems discussed earlier exist in a particular community or for a particular physician's patients, the more likely it is that a spending measure will change from year to year, since the measure will be reflecting the influence of many factors other than how the individual provider is delivering care. Assessing this requires a different measure of reliability – *temporal reliability* or *test-retest reliability*. Even less research has been done on temporal reliability of spending measures; this is particularly challenging to study, because it is difficult or impossible to distinguish changes in spending resulting from an intentional change in a provider's behavior from changes in spending due to changes in patient characteristics or changes in the treatment patterns of other providers who are involved in the patient's care. However, the research that does exist suggests that one should be very cautious about using current spending measures as predictors of future performance. For example:

- The study by Yu and colleagues cited earlier found that for most types of utilization measures examined for primary care physicians and practices, the test-retest reliability of the measures was much lower than the cross-sectional reliability.[69]
- Although a reliability analysis published by HealthPartners to support its Total Cost of Care measure primarily focused on cross-sectional reliability, it also included data on how the measure varied over time for 19 provider groups.[70] The study said that providers' performance was "relatively consistent across all three years," but the data presented showed that the average year-to-year variation in spending for individual providers was more than double the average difference *between* providers in any year, and as a result, the relative cost rankings of some providers changed significantly within two years.[71] A companion analysis of the HealthPartners Resource Use measure showed that the average annual variation in resource use for individual providers was more than triple the average difference between providers in any year, and the provider ranked as the fourth lowest in resource use in one year was ranked as the provider with the highest resource use two years later.[72]
- A study done for CMS of the stability of physician scores based on composite episode-based spending measures of services to Medicare beneficiaries found that the one-year correlation of physician scores was at most .60, that physicians classified as being the highest-cost providers in a given year had less than a 50% likelihood of being classified the same way in the following year, and only 40% of providers remained in the same score category over a two-year period.[73]
- A study that examined spending patterns for patients assigned to Accountable Care Organizations using standard attribution methodologies found that only two-thirds of patients who could be assigned to an ACO in either of two years were assigned to the same ACO in both years. Patients in the highest decile of spending were the least likely to be assigned to the same ACO in both years, and they also were the most likely to use PCPs outside of the ACO during the year, meaning that an ACO's spending could change from year to year simply because of changes in the number of high-need/high-utilization patients who are assigned to it in each year, even if there were no changes in the underlying approach to care by the individual providers in the ACO.[74]

Indeed, since research has shown that there is significant regression to the mean in patient spending (i.e., most of the patients who received high levels of expensive services in one year will not have similarly high spending the following year), it would not be surprising to find that a provider who had above average per-patient spending in one year would have lower per-patient spending in the following year; the high per-patient spending in the first year may simply have been because a subset of their patients had unusually high needs for expensive services that were not adequately captured by risk adjustment systems due to the many limitations described earlier.

## Variation in Reliability Within and Across Communities

It is important to recognize that the "reliability" of a spending measure is not just a function of the methodology used, but it will vary depending on the specific physician practices being measured, the types of patients they care for, and the characteristics of the community where they practice, even if the same methodology is used. In fact, the limitations of attribution and risk adjustment methodologies described earlier can make a spending measure unreliable in ways that can only be identified after the fact. If the difference in per patient spending between two providers is due to unmeasured or unadjusted differences in the needs of their patients rather than to actual

differences in the way the providers deliver care to similar patients, then the spending measures might be very unreliable indicators of true differences in the providers' performance and very poor predictors of the actual spending that would occur if those providers were to begin seeing different types of patients. Similarly, if different subsets of patients are attributed to a provider each year (even though the size of the provider's total patient panel is unchanged), then spending measures may change over time even though the provider's underlying approach to care has not changed.

*EXAMPLE: Assume that there are two groups of patients in the community, each with the same chronic disease. The individuals in Patient Group 1 visit their primary care practice regularly, reliably use medications to manage their chronic disease effectively, and only rarely visit the emergency room or have to be hospitalized for exacerbations of their chronic disease. The individuals in Patient Group 2 also visit their primary care practice regularly, but they do not reliably manage their chronic disease effectively and end up going to the emergency room and being admitted to the hospital every year for an exacerbation of their chronic disease. Both groups of patients have the same risk score based on their chronic disease and other characteristics measured by the risk adjustment methodology, but as shown in Figure 8, the average total spending on patients in Group 2 is much higher than on the patients in Group 1 because of the frequent hospitalizations in Group 2. Assume further that there are two physician practices in the community; each practice manages the care of each group of patients identically, so the spending per patient within each group of patients is identical in each practice. However, the mix of patients from the two groups differs significantly between the two physician practices. 80% of the patients in Physician Practice 1 are from Patient Group 1, whereas 80% of the patients in Physician Practice 2 are from Patient Group 2. As shown in Figure 8, the average spending per patient in Physician Practice 2 is nearly three times as high as in Physician Practice 1, even though there is no difference in the way the physicians treat the patients, simply because of the different mix of patients in each practice.*

*Based on the spending measure, one of the health plans in the community determines that Physician Practice 2 is inefficient in managing chronic disease patients and removes Physician Practice 2 from its network. As a result, half of Physician Practice 2's chronic disease patients move to Physician Practice 1, and these patients are predominantly from Patient Group 2. Because Practice 1 is now much busier with the influx of the patients that had previously been using Practice 2, some of the Group 1 patients from Practice 1 move to Practice 2. Now, a majority of the patients in Physician Practice 1 are from Patient Group 2, and a majority of patients in Physician Practice 2 are from Patient Group 1. As a result, the average spending per patient in Physician Practice 1 is now 37% higher than in Physician Practice 2, even though both practices have continued to deliver care in exactly the same fashion as in the previous year and each practice delivers care in exactly the same way that the other does.*

## Conditions Supporting Greater Reliability and Validity of Measures

The limited research available suggests that the significance of the problems will vary depending on the community where spending measures are being created, the exact methodology used in calculating the measures, and the purposes for which the measures are being used. Ideally, sensitivity analyses would be conducted to determine how serious the problems are in the particular community where spending measures are being generated and to assess the impacts of those problems on the purposes for which the measures will used.

In the absence of such sensitivity analyses, some qualitative judgments can be made about the conditions in which current attribution and risk adjustment methodologies will provide more valid, reliable, and actionable comparisons of providers on spending.

- **Organizational Structure of Providers.** If most physicians in the community are part of a large multi-specialty group or a health system, then even if the individual physician to whom the patient is assigned did not *directly* deliver or order all of the services that a patient receives, it will be far more likely that most or all of those services are controlled by *some* provider who is part of the group or health system. If the spending measure is reported or used at the group or system level, rather than at the individual physician level, differences in spending will be far more likely to reflect differences in the collective approach to care delivery used by the providers in that group or system.

  Moreover, in communities where most physicians work for a large health system, many of the problems of the unattributed patients described earlier can be solved simply by assigning patients to the health system based on all of the services the patient receives, rather than just on primary care visits. For example, if a patient is a frequent emergency room utilizer and does not see a primary care physician at all, the health system that operates the emergency room is, as a practical matter, providing the patient's primary care and could be assigned the spending for that patient. Moreover, if the patient were assigned to the health system based on emergency room utilization as well as on physician office visits, there would be no penalty in terms of the spending measure if a primary care physician affiliated with the health system began seeing such a patient, because the patient's spending would be assigned to the health system either way. In order to avoid excluding relatively healthy patients, patients who had not made an office visit to a primary care physician recently could be assigned to the health system where the patient had received the majority of all types of care or where the patient had seen a primary care physician most recently, even if it was several years earlier.

  In contrast, in communities where significant numbers of physicians are in small or specialty-specific

## FIGURE 8

### Change in Spending Rankings for Physician Practices Resulting from Changes in Patient Case Mix

Utilization and Cost in Both Practices

| | Services Per Patient Per Year | Spending Per Service | Spending Per Patient Per Year |
|---|---|---|---|
| **Patient Group 1** | | | |
| PCP Visits | 3.0 | $100 | $300 |
| ER Visits | 0.1 | $750 | $75 |
| Hospital Admissions | 0.1 | $10,000 | $1,000 |
| Total Spending | | | $1,375 |
| **Patient Group 2** | | | |
| PCP Visits | 3.0 | $100 | $300 |
| ER Visits | 1.0 | $750 | $750 |
| Hospital Admissions | 1.0 | $10,000 | $10,000 |
| Total Spending | | | $11,050 |

| | YEAR 1 | | | | YEAR 2 | | | |
|---|---|---|---|---|---|---|---|---|
| | Physician Practice 1 | | Physician Practice 2 | | Physician Practice 1 | | Physician Practice 2 | |
| | # of Patients | Spending | # of Patients | Spending | # of Patients | Spending | # of Patients | Spending |
| Patients in Group 1 | 800 | $1,100,000 | 200 | $275,000 | 400 | $550,000 | 600 | $825,000 |
| Patients in Group 2 | 200 | $2,210,000 | 800 | $8,840,000 | 600 | $6,630,000 | 400 | $4,420,000 |
| Total Patients | 1000 | $3,310,000 | 1000 | $9,115,000 | 1000 | $7,180,000 | 1000 | $5,245,000 |
| Spending Classification | Low | | High | | High | | Low | |
| Spending Per Patient | $3,310 | | $9,115 | | $7,180 | | $5,245 | |

practices or when measures are being applied to individual physicians, spending measures using current attribution methods are less likely to produce meaningful and actionable comparisons of physicians.

- **Availability of Multiple Providers.** Accountability systems that attribute total spending for patients to primary care physicians (PCPs) implicitly assume that PCPs can control all or most aspects of spending through their decisions about ordering tests and referrals to specialists. However, even though primary care physicians may *want* to control spending, their *ability* to do so will depend on the choices available to them and their patients. In many geographic areas, there may be only one specialist, laboratory, hospital, or other provider that delivers a particular type of service or is accessible to the patient. Consequently, comparing total spending for PCPs across communities will be more problematic than comparing PCPs within the same community.
- **Patients' Health Plan Benefit Design.** If a spending measure is based solely or primarily on patients who have a health plan benefit design that requires them to designate a primary care physician or if there are strong incentives in the patients' health plan for them to use specialists recommended by their primary care physician, then it is far more likely that a measure of the total spending on the patients would reflect services that were delivered or ordered by the primary care physician than if the spending measure is based primarily on patients who had the ability to independently choose and use specialists for care without involvement of the primary care physician.
- **Homogeneity and Stability of Patient Populations.** If spending measures are being generated for patients who are relatively similar in age, health, income, education, functional ability, etc. and if the same patients are being measured over time (e.g., the workers of an employer with relatively little turnover or the residents of a community with relatively low in-migration), then the weaknesses of common risk adjustment tools will be less of a concern than if the spending measures are being applied to a very heterogeneous population or a population that is changing significantly over time.
- **Participation in Primary Care Medical Home Support Programs.** If the primary care providers in the community are receiving additional resources to help them manage patient care in return for accountability for managing total spending for patients who are assigned to them, then regardless of the flaws in the attribution methodology, the providers will have had greater opportunity to make changes either in the services they delivered or the patients they accepted in order to improve their performance on those measures than providers who are merely being measured retrospectively with no additional resources to improve care.
- **Purpose of Measurement.** If the purpose of generating spending measures is to help communities identify where there are opportunities to reduce spending, not to assign accountability for spending to specific providers, then the issues associated with attribution disappear. For example, county by county comparisons of spending could show that the residents of one county use emergency rooms at a much higher rate than other counties and use primary care practices much less often than other counties, which could reveal that a shortage of primary care physicians is causing overall healthcare expenditures to be high. The problem would not be solved by blaming current primary care physicians for the high spending, but by recruiting additional primary care physicians to the community.

  If attributing spending measures to providers is intended solely to help providers identify potential opportunities for reducing spending, then the lack of precision in the attribution methodologies as well as the limitations in risk adjustment and the other problems identified earlier are of less concern, since the spending measure will presumably only be the first step in a multi-step process of analysis and action. For example, if an attributed spending measure indicates that a particular provider has high spending, further analysis could be done to determine the reasons for that, and if the causes were determined to be factors outside of that provider's control, attention could then be redirected to those factors. Multi-stakeholder Regional Health Improvement Collaboratives are often in the best position to carry out these kinds of additional analyses.

  In contrast, if spending measures are used by payers to change payment to providers or to define provider networks, the problems described in this section will be of serious concern. If the spending measure is converted directly into a payment bonus or penalty, then there is no opportunity to do further analysis to determine which factors caused differences in spending and whether those factors were within the provider's control, and a provider could be unfairly penalized or rewarded. Moreover, the payment structure could end up rewarding providers for avoiding high-cost patients instead of encouraging them to find ways to reduce spending on those patients. Spending measures based on services delivered to specific groups of patients in the past may not be predictive of how the same providers would deliver services to different groups of patients in the future, yet payers are implicitly asserting that the measures are predictive when they define narrow networks using such measures and steer new groups of patients to those providers.

Ideally, providers in a community should participate in the development of the methodology that will be used to generate spending measures so they can determine (a) whether the problems identified in this section will have a large or small impact given the structure of providers, payment systems, and benefit designs in the community and (b) whether the measures will be used in a way that avoids any undesirable impacts on the quality of care in the community. Moreover, participation by providers can help to develop better measures of accountability, as described in Section IV.

# IV. BETTER WAYS OF MEASURING AND ASSIGNING ACCOUNTABILITY FOR HEALTHCARE SPENDING

The serious problems identified in Section III do not mean that efforts to measure spending and assign accountability to providers should be abandoned. For too long, efforts to control healthcare costs have focused narrowly on controlling the prices of individual services; the result of cuts in fees and demands for large discounts in prices has not been lower spending but faster growth in utilization of services. Recent efforts to promote "transparency" about prices could have a similar effect, since a provider could be rewarded for having lower prices on *individual* services even though that provider uses *more services* and causes *more complications* for patients than other "more expensive" providers.

The solution is develop better measures of spending and better methods of assigning accountability for spending that explicitly address the issues described in Section III. To be maximally effective, methods for measuring spending and assigning accountability should have the following five capabilities:

A. Identifying the services and spending that can be controlled or influenced by each provider;

B. Identifying which services represent opportunities for reducing spending without harming patients;

C. Determining which patients have greater needs for services;

D. Adjusting for structural differences in costs for different providers; and

E. Comparing providers based on both costs and outcomes of care.

These components are described in the sections that follow.

## A. Identifying the Services and Spending Providers Can Influence

The first step in more effectively identifying ways of reducing healthcare spending without rationing and identifying the healthcare providers best able to make the reductions is to divide spending into categories that reflect differing levels of provider control or influence over services. Five such categories are:

1. Services both *ordered and delivered* directly by the physician or other provider who is being measured.
2. Services delivered by *other* providers that are *integrally related* to services delivered by the provider being measured.
3. Services delivered by *other* providers that resulted from *orders* or *referrals* from the provider being measured, and services delivered by the provider being measured in response to an order from another provider.
4. Services delivered by other providers that were *related* to services delivered or ordered by the provider being measured.
5. All other services the patient received that are *unrelated* to services delivered or ordered by the provider being measured.

Collectively, these five categories add up to the total spending on a patient from all providers, but the categories are more understandable and actionable for individual providers than traditional actuarial breakdowns. The services and spending included in each category will differ for different types of physicians and other providers, since different providers deliver and order different kinds of services. However, in all cases, a provider will have greater influence over the lower-numbered categories than the higher-numbered categories, so this categorization will better identify which providers could actually reduce spending than current attribution methodologies which simply attribute the spending in all five categories to a single provider who happened to provide a certain proportion of the overall services. Moreover, under this approach, *every provider* will have *all* of the spending *they directly control* attributed to *them*, rather than attributing it to a different provider who happened to deliver more services to the same patients.

Within each of these categories, it will be helpful to further disaggregate spending in order to distinguish which services represent greater opportunities for spending reduction; methods for doing this will be described in Section IV-B.

### Spending Category 1: Services Ordered and Delivered by a Provider

Spending Category 1 consists of services that are both ordered and delivered by the physician, hospital, or other provider for whom spending is being measured. (Category 1 would exclude services the provider delivered in response to an order from another provider; these services are discussed in conjunction with Spending Category 3 below.) The services in Category 1 are those for which the provider has the most direct control over costs and also quality. If the physician or other provider decides to deliver unnecessary services, delivers services inefficiently (e.g., conducting multiple tests when one would be sufficient), or charges a high price for individual services, that provider has the ability to directly lower spending on services for the patient by reducing the number of services delivered, reducing the prices charged for individual services, or both.

For many physicians, all or most of their services are ordered and delivered by them in their offices, and these services represent most or all of the spending they can

control. These services can be readily identified from the claims filed or payments made to each provider.[75] In some cases, these services involve significant amounts of spending.[76]

> ***EXAMPLE:*** *Ophthalmologists administer very expensive drugs in their own offices to treat patients with macular degeneration and other eye conditions. They decide whether treatment is appropriate and what drug to use, and their decisions can have significant impacts on spending.[77] Most patients see an ophthalmologist with no involvement by a primary care physician or other doctor, so it does not make sense to "attribute" the ophthalmologist's services to a primary care physician or any other physician other than the ophthalmologist.*

Specifically identifying the services and spending in Category 1 can help to control healthcare spending. There are many examples of how savings have been achieved in healthcare merely by measuring the types of services that physicians directly perform and showing them that other physicians deliver fewer or different services with equal or better outcomes for their patients.[78]

### *Including Patients With No (Paid) Services*

As explained in Section III-A, one of the problems with claims-based attribution systems is that patients who received no services at all within the relevant lookback period are not attributed to any physician. The fact that they received no services for which a *claim was filed* does not mean that they received no services *at all* or that no physician should be accountable for the patient. Some patients' health issues can be effectively managed by a physician without the need for any paid service claims. For example, a patient with hypertension may be managing their condition effectively with diet and exercise recommended by the physician, and the patient may be taking blood pressure readings at home and reporting them to the physician without the need for frequent office visits, but for these patients, there would be no spending to be assigned to the physician. Some physicians encourage patients to come to their offices for blood pressure checks at no charge so that visit copayments do not deter them from regular blood pressure monitoring, but if no claim is filed for the office visit in order to avoid the need to charge a copayment, there would also be no spending in a claims file to enable the patient to be attributed to that physician.

However, since spending measures for physicians are calculated and compared on a per-patient basis, if these well-managed patients are not counted in the denominator when the measure of spending is calculated for a physician, the physician's per-patient spending will be misleadingly high. All else being equal, the more patients a physician has whose care is managed effectively without an office visit or other billable service, the worse the physician will look on spending measures compared to other physicians based on the patients they do see. As noted in Section III-A, this problem can be exacerbated in patient-centered medical home programs and other programs that make payments to physicians on a per-patient, rather than a per-service basis. If the physician practice receives funding that it can use to begin delivering otherwise unreimbursed services such as making telephone calls and sending emails to patients, using nurses to make home visits, etc., patient outcomes can improve while billable services decline, but this can also mean that per-patient spending measures based solely on billable services will increase.

This issue is also important for adequately addressing services associated with potentially preventable conditions (these are discussed in more detail in Section IV-B). If a patient made frequent visits to the emergency room not because they had no primary care physician, but because their primary care physician was not accessible and the patient was forced to use the emergency room instead of seeing the PCP, the emergency room service should not only be assigned to the emergency room physician and hospital (because they delivered that service), but also to the primary care physician (the emergency room visit should be included in Spending Category 4 for the PCP, as described in more detail below). However, one needs to know that the patient *had* a PCP in order to attribute the ER visit to that PCP.

The ideal solution to this would be to have patients explicitly designate a particular physician as their primary care physician or designate a particular specialist as the physician managing their care for a specific condition. If these patients did not receive any billable services from that physician during the time period being measured, then the patient could still be included in the *denominator* of per-patient spending measures for that physician, even though there is no spending for them in the *numerator* in Spending Category 1. This type of designation does not mean that the patient must be "locked in" to the primary care physician and specialists they designate, or that the primary care provider the patient designates has to serve as a gatekeeper and give his or her approval before the patient can see a specialist. These restrictions are used in HMO health plans to control spending, but in this case, all that is needed is an indication that a patient views the physician as managing their care at a particular point in time. A patient could change the designation at any time, but while the designation of a physician was in effect, that patient would be counted as part of the physician's patient panel whether they received billable services or not.

In the absence of such designations, the information in Spending Category 3 (described below) can help reduce the number of patients that are inappropriately unassigned. If a physician *orders* a service for a patient from another provider, even if the physician did not *deliver* a billable service to the patient, then it can be assumed that the physician is playing a role in managing the patient's care for the condition(s) for which the services were ordered. It will be particularly important to have access to pharmacy claims data here; many relatively healthy patients whose care is being managed by a physician will be taking medications prescribed by that physician but receive no other billable services from them, since the patient may simply call in to the physician's office for prescription refills and only make infrequent in-person visits to the physician's office.

An alternative approach will become increasingly feasible as Medicare and other payers begin calculating quality measures based on data collected from electronic health

records. If a physician practice includes a patient in the denominator for a quality measure reported from an EHR, then it would make sense to include that same patient in the denominator for spending measures for the same practice. This could significantly increase the number of patients assigned to physicians; one study found that only half of the patients on whom primary care physicians reported quality measures to Medicare would have had those patients attributed to them based on receiving a plurality of claims-based evaluation and management visits.[79]

## Spending Category 2: Services Integrally Related to the Provider's Services

In many cases, a service in Category 1 is routinely delivered in conjunction with a service that a different physician, a hospital, or another provider delivers. Under fee-for-service payment, these are treated as separate services and the two providers are paid separately. However, from the patient's perspective, the two providers jointly delivered a single "bundle" of services in order to address the patient's need. These integrally related services would be included in Spending Category 2.

*EXAMPLE: If a physician performs a procedure at a hospital, the physician and hospital will be paid separately, but the hospital could not have performed the procedure without the physician, and the physician may not have been able to perform the procedure without the hospital, so the two services are integrally related. The physician's payment should be included under Spending Category 1 in a spending analysis focused on that physician; the hospital's payment would then be included under Spending Category 2 for that physician. Similarly, the hospital's payment should be included under Spending Category 1 in a spending analysis focused on the hospital, and the physician's payment would then be included in Spending Category 2 for that hospital.*

The patient (or their third-party payer) will have to pay for all of the components when the patient receives the bundle of services, but different sets of providers may charge different amounts for the different components or use different combinations of components to deliver the same overall service bundle. It is impossible for the patient to accurately compare two providers' prices or spending for a service unless all of the components are included. Many patients have experienced receiving multiple bills from different providers for different parts of the "same" service and being surprised to find that the total was far more than they expected. Including the spending on all of the services in a spending analysis for each of the providers involved more accurately reflects the total cost of the services they are delivering. However, since the provider being measured will have greater control over the cost of the components they deliver directly than over the components delivered by the other providers, it makes sense to keep them in separate categories.

*EXAMPLE: If a gastroenterologist performs a colonoscopy in her own office, a single amount will be paid to the gastroenterologist to cover the entire cost of the procedure, including both the gastroenterologist's time and the "facility" costs of the office where the procedure is done. However, if the gastroenterologist performs the colonoscopy as an outpatient procedure in a hospital, one payment will be made to the gastroenterologist, and a separate payment will be made to the hospital. The hospital payment would not have been made if the gastroenterologist had not performed the colonoscopy at the hospital, and so the two payments should be considered together as a "bundle," rather than as independent services by the physician and hospital.[80] If the two payments are not examined together, it would appear that the gastroenterologist performing a colonoscopy in the office is more expensive (since the gastroenterologist is paid more for a procedure done in the office than in the hospital), even though, in general, the total cost for a colonoscopy performed in a hospital is higher after combining the lower payment to the gastroenterologist with the separate payment made to the hospital. The hospital payment should be shown in Category 2 because the gastroenterologist can control whether the procedure is done in the hospital but may not be able to control what the hospital is paid when the procedure is done there.*

*Similarly, if a gastroenterologist chooses to use a form of sedation on the patient that requires the presence of an anesthesiologist who bills separately for that service, the payment to the anesthesiologist should be included in Spending Category 2 for the gastroenterologist, since the anesthesiology payment would not have occurred if the gastroenterologist had not performed the colonoscopy. Since the gastroenterologist chose to use a type of anesthesia requiring an anesthesiologist, the gastroenterologist should be accountable for the services billed by that other provider.*

Despite the obvious connection between physician spending and hospital spending, the federal government has failed to incorporate this linkage into its policies regarding physician payment for over two decades. The Sustainable Growth Rate formula requires cuts in physician fees if total spending on *physician services* increases faster than growth in the economy as a whole, regardless of whether *total spending* on *all services* is increasing as quickly. For example, if physicians provided more services to patients that helped them stay healthier and avoid hospital stays, the Sustainable Growth Rate formula could force cuts in the physicians' fees (because spending on physician services increased) even though total spending (including hospital stays and other Medicare Part A services) decreased.

### *Identifying "Bundles" of Related Services*

Because each provider is paid separately under fee for service payment systems, one has to look for all of the individual claims from all of the different providers who were involved with a particular treatment or procedure to identify those providers and services that should be assigned to Spending Category 2; no single claim form will have all of that information.[81] There is not a standard way today of identifying and combining these "bundles" of related services. The episode groupers discussed in Section

ll generally group a much larger collection of services over a longer period of time than what is appropriate for Spending Category 2 as defined here (these other services will be discussed separately below), and in many cases, the groupers may not group the services that should logically be bundled.[82]

The vast majority of services and procedures patients receive are delivered in outpatient settings.[83] In these cases, the integrally related services for outpatient care that should be included in Category 2 will typically occur on the same day and they will have the same or similar billing codes, at least for the principal physician and the hospital services.[84] Definitions for some of these kinds of bundles have been developed,[85] and more such definitions are needed.

The process of identifying integrally related services is more difficult when a patient is receiving multiple procedures or services on the same day, since it is more likely that multiple providers will be performing separate services that are not integrally related. Identifying bundles of services is even more difficult if multiple services were delivered during an inpatient hospital stay, since the hospital will typically not be paid separately for the individual services,[86] and since different coding systems are used for classifying physician services and inpatient hospital services.[87] In these cases, two options could be used for determining which services to include in Spending Category 2 for a particular provider:

- The most conservative option would be to include a service from another provider that occurs on the same day but only if it has a service/procedure code that is the same as or similar to the code billed by the provider in Spending Category 1. This approach would miss some services from providers that bill for services in different ways (e.g., it would likely miss the anesthesiologist's payment as part of the colonoscopy in the earlier example), but it would avoid including completely unrelated services that happened to occur on the same day.
- A less conservative option would be to include *any* service that occurs on the same day as a service included in Spending Category 1. This would likely include some unrelated services, particularly for patients with multiple health problems who may receive multiple services at the same time, but this approach would be more likely to include all integrally related services that are delivered in an outpatient setting. The payment for an inpatient hospitalization could be divided by the length of the stay in order to calculate per diem amounts and the per diem amount could then be assigned to any physician who billed for a service on a day during that stay.

The ideal is for teams of providers to accept *prospectively*-defined *bundled payments* for these combinations of services for patients who have chosen those providers to deliver the services. As discussed in Section VI, retrospectively analyzing spending in coherent bundles will help facilitate the creation of prospectively defined bundled payments.

### *Distinguishing Price Level vs. Bundle Efficiency*

As noted earlier, it is important to track different providers' components of a bundle separately in Spending Category 1 and Spending Category 2 because of the difference in the ability of the provider to control the spending in the two categories.

Clearly, if more individual billable services or more expensive services are used to deliver the total bundle associated with what is a single service from the patient's perspective, that bundle will be more expensive. In the Medicare program, the prices of individual services are defined by federal regulation, and so the principal differences in the total cost of services across providers in Medicare will result from differences in the services included in the bundle, i.e., the "efficiency" of the bundle.[88]

However, for patients who have commercial insurance, the prices of the same service or bundle of services can vary dramatically from provider to provider. In the colonoscopy example, one gastroenterologist's colonoscopies could more expensive than another's because the hospital where the colonoscopies are done charges higher prices, even though all other aspects of the colonoscopy are the same (i.e., both physicians do the colonoscopies in a hospital instead of their office and neither use an anesthesiologist to administer sedation). If the patient needs to have the colonoscopy done in a hospital and if there is only one hospital available, then the physician may have little or no control over the biggest portion of the total spending, but if the gastroenterologist could do the procedure outside of the hospital, the physician would have much more control of the total cost.

To distinguish pricing vs. efficiency, spending on the services in Category 2 can be divided into three factors: (1) the number of components used as part of the overall service, (2) the "reasonable" unit cost or payment for each component, and (3) the ratio of the actual unit price of the service components to the reasonable cost of those components. Mathematically, this can be represented as follows:

Number of Service Components Provided in Service Bundle
x
Reasonable Cost Per Component
x
Ratio of Actual Payment to Reasonable Cost
=
Total Spending on Service Bundle

The first factor measures the provider's decision to use multiple components in the overall service bundle, the second factor measures how expensive those components would be under the *best* of circumstances, and the third factor measures how expensive the service components *actually* were from the providers that delivered them. If total spending on a particular service bundle is high for a physician's patients relative to other physicians' patients, then this breakdown allows the physician to identify the most effective ways to reduce that cost:

- if unnecessary service components are being delivered as part of the bundle, the physician managing the overall service bundle could reduce spending by eliminating the unnecessary components;

- if the service components being used are inherently expensive and there are lower-cost options available, the physician could reduce spending by using the lower-cost options;
- if the ratio of actual payments to reasonable cost is high, the physician could determine whether there are lower-priced providers available who could be used to deliver the same service components at a lower price.

*EXAMPLE: An orthopedic surgeon performs knee replacements at a particular hospital. The surgeon could reduce the cost of the knee surgery by using a lower-cost knee implant or by performing the surgery in a lower-cost hospital.*

There are different ways that what is referred to here as the "Reasonable Cost Per Component" factor could be calculated. One approach would be to define it as the payment rate that Medicare uses for the service.[89] An alternative approach would be to take the average amount that commercial payers pay for the services across multiple providers; these amounts would be more consistent with the payments actually made for services delivered to commercially insured patients than Medicare payment amounts, but the amounts would likely only be "reasonable" if the local market were truly competitive. A third approach would be to calculate the relative amounts of resources typically used in delivering different service components.[90] The ideal approach would be to determine the lowest achievable cost to deliver a particular service in the local market, since payment rates and even relative resource measures may bear little relationship to the costs that would exist under efficient service delivery. However, information on the actual costs of services and how those costs would change in response to changes in volume and changes in the method of care delivery are rarely available.[91]

Figures 9 and 10 show a hypothetical example of what a report that included this type of information might look like for gastroenterologists performing colonoscopies. The report uses payment amounts similar to what are paid by Medicare as an indicator of reasonable payment amounts for the services. Figure 10, which is based on the data in Figure 9, makes it clear that even though the *fees* charged by the first group of gastroenterologists are much higher relative to Medicare payment rates than either of the other two groups, their colonoscopies still cost 20% - 30% less *in total* than the others because they do the procedures in the office, rather than in the hospital.[92] Conversely, although the physicians in Practice 2 have the lowest fees for their own services, their colonoscopies are more expensive *overall* than either of the other groups because they perform the procedures at a hospital with relatively high prices relative to Medicare payment levels. From a patient or payer's perspective, the fact that the gastroenterologists in Group 1 charge much higher prices relative to Medicare should be less of a concern given that the total package of services they deliver is less expensive.[93]

This demonstrates that "transparency" about prices on individual services from individual providers is not enough to help consumers identify the most cost-effective providers, and in some cases it could lead them to make the wrong choices. Consumers need to know the *total* combined cost of *all* of the components that will be in the total service bundle they receive from all of the providers involved. In the example above, a consumer who chose Practice 2 because the gastroenterologists there charged the lowest physician fees among the three practices would find that they actually paid more in total to get a colonoscopy from Practice 2 than if they had they used either of the other practices, and they paid 46% more than if they used Practice 1 which appeared to be the most expensive based on physician fees alone.

## Spending Category 3: Services Delivered Based on Orders or Referrals

### *Services Resulting from Orders by a Different Provider*

Spending Categories 1 and 2 focused on the services that were *ordered by and delivered to* a patient by the provider who is being measured and the *integrally related* services from other providers. However, a significant amount of healthcare spending is associated with services that one physician or other provider *orders* for the patient but a different physician or provider actually *delivers* and is paid separately for. For example, if a primary care physician orders an X-ray, MRI, or other imaging study for a patient, the imaging study will often be performed by an imaging facility that is independent of the primary care practice, and the interpretation of the image will be done by a radiologist or other specialist from a different physician group. Both the imaging facility and the physician interpreting the image will be paid separately from the physician who ordered the imaging study. Similarly, medications are ordered (prescribed) by physicians, but except for medications administered in a hospital or in a physician's office, most patients will receive the medications from a pharmacy and the medications will be paid for by the patient or their drug insurance through the pharmacy, not through the physician who prescribed the drugs.

If spending measures and accountability were focused solely on what is included in Spending Categories 1 and 2, there could be an incentive for one provider to order a service from another provider rather than to deliver the service themselves. Consequently, it is important to include the spending on services a provider orders in an analysis of the spending they can control or influence.

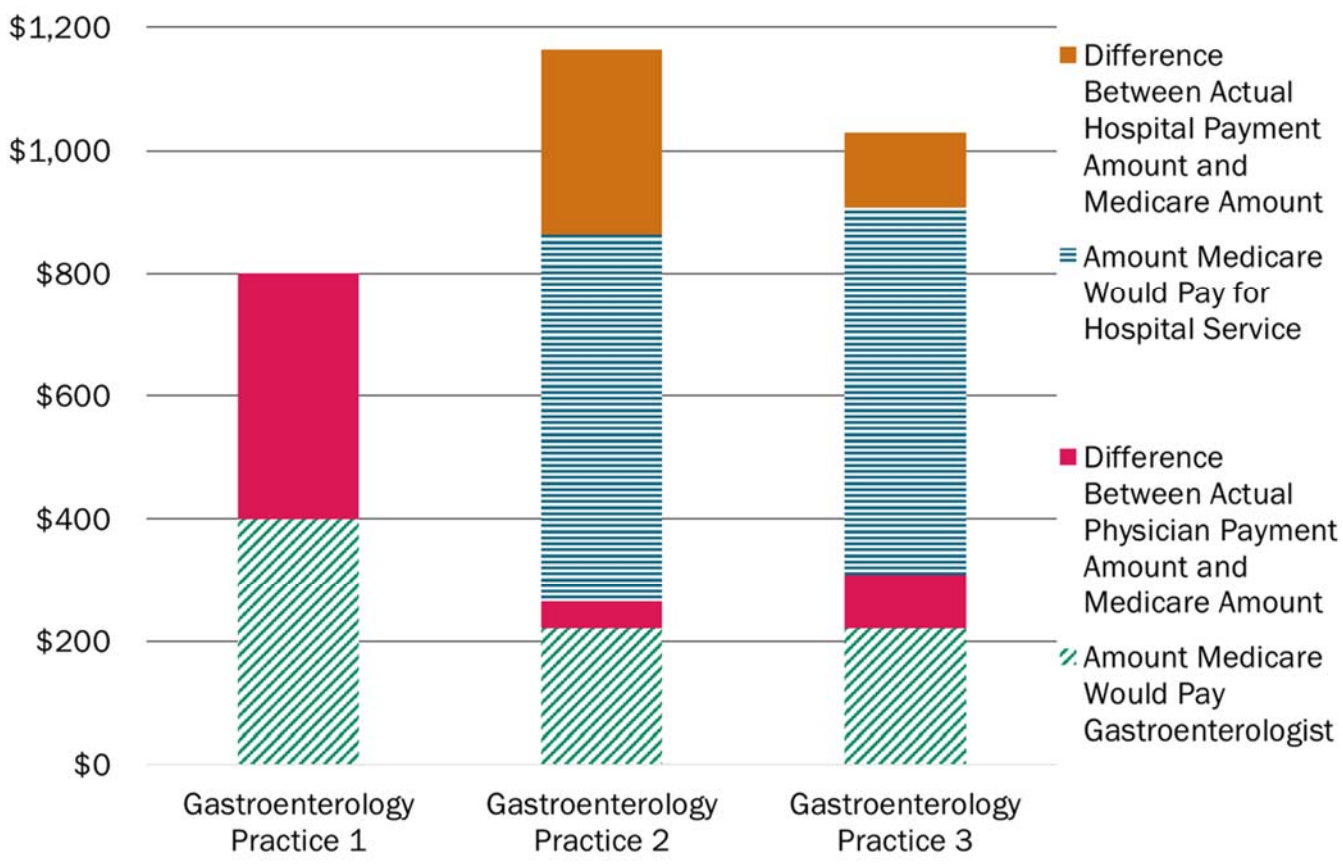
However, even though the provider who ordered the service should clearly be accountable for *making* the order, it is usually not appropriate to assign them full accountability for the *total spending* that *results* from that order. Although the spending on the service would likely not have occurred if the ordering provider had not ordered it, once the service was ordered, the ordering provider likely would not have any control over exactly how the service would be delivered or how much would be paid for it, whereas the provider who delivered the service may well have choices about how to deliver the service that would affect its costs. In other words, there is *shared* accountability between the provider who *ordered* the service and the provider who *delivered* the service.

These services should be included in Spending Category 3. However, because there are distinct but inter-related

## FIGURE 9
### Differences in Commercial Insurance Spending on Colonoscopies at Three Hypothetical Gastroenterology Practices

| | Gastroenterology Practice 1 | Gastroenterology Practice 2 | Gastroenterology Practice 3 |
|---|---|---|---|
| **Category 1: Gastroenterologist Services** | | | |
| Medicare Payment Amount | $400 | $220 | $220 |
| Ratio of Actual to Medicare Payment | 200% | 120% | 140% |
| Actual Payment To Gastroenterologist | $800 | $264 | $308 |
| **Category 2: Hospital Services** | | | |
| Medicare Payment Amount | | $600 | $600 |
| Ratio of Actual to Medicare Payment | | 150% | 120% |
| Actual Payment to Hospital | | $900 | $720 |
| **Total Spending Per Colonoscopy** | $800 | $1,164 | $1,028 |

## FIGURE 10
### Components of Spending for Colonoscopies in Different Gastroenterology Practices



roles played by the two providers, instead of simply assigning the *total spending* on these ordered services to one provider or the other or both, it is more appropriate to assign *shares* of the spending to each provider based on each provider's relative influence over the spending.

- Since the ordering physician controls whether the service is ordered and the number of services ordered, accountability for the *utilization* of a service should be assigned to the *ordering* physician.
- In addition, the ordering physician also should be accountable for whether a type of service was ordered that would, in general, be more expensive (even if delivered by the most efficient providers) than an alternative service that could also have been ordered for the same patient problem (e.g., a less extensive blood test). However, the ordering physician should not be held accountable for whether the delivering provider was paid more or less than would be expected for the service (which in turn would depend on both the number of components that provider used to deliver the service and the prices charged by that provider for each component).
- The delivering provider should then be accountable for the amount by which the actual spending per service differs from what would be expected from an efficient provider for the services that were ordered.

In mathematical terms, the *ordering* provider is responsible for the following subset of the spending:

Number of Services Ordered Per Patient
x
Reasonable Spending Per Service
=
Expected Spending Per Patient

The provider who actually *delivers* the service is responsible for the following subset of the spending:

Actual Payment Per Patient
Minus
Expected Spending Per Patient
=
Delivery Payment Differential Per Patient

The sum of these two amounts is equal to the total spending per patient on this type of service.

*EXAMPLE: In one community, gastroenterologists perform colonoscopies in the office, and are paid in a fashion similar to Gastroenterology Practice 1 in Figure 9. In a second community, gastroenterologists perform colonoscopies in the hospital, and are paid similar to Gastroenterology Practice 2 in Figure 9. In each community, a primary care practice orders colonoscopies for 200 patients. If one assumes that a payment of $480 for an office-based colonoscopy is a reasonable amount, Figure 11 shows that in each case, the primary care physician would be assigned $96,000 in Spending Category 3, and the remaining spending would be assigned to Spending Category 3 for the gastroenterology practice. The colonoscopy spending per practice in Community 2 is nearly 50% higher than in Community 1, not because PCPs are ordering more colonoscopies, but because the gastroenterologists in Community 2 spend 50% more to deliver a colonoscopy.*

Some people might argue that the ordering physician should be held accountable for the *total* actual spending per patient, rather than splitting the spending into these two components. However, in traditional Medicare or a PPO health insurance plan, the patient chooses where to get the service that the physician ordered, not the physician. Moreover, in many geographic areas, there may be only one provider who delivers a particular service or is accessible to the patient. Even if there are multiple pro-

## FIGURE 11
### Colonoscopies Ordered by Primary Care Practices in Two Hypothetical Communities

| | Community 1 | | Community 2 | |
|---|---|---|---|---|
| | Spending Assigned to Primary Care Practice | Spending Assigned to Gastroenterology Practice | Spending Assigned to Primary Care Practice | Spending Assigned to Gastroenterology Practice |
| **Category 3: Ordered Services** | | | | |
| Assumed Reasonable Payment of $480 Per Colonoscopy for 200 Colonoscopies | $96,000 | | $96,000 | |
| Actual Payment for Colonoscopies Above $480 | | $64,000 | | $136,800 |
| **Total Payment for 200 Colonoscopies** | $160,000 | | $232,800 | |

viders available, physicians typically have little or no knowledge of the actual amounts paid to different providers for their services, particularly if those providers are part of different organizations. Even in narrow network health plans, the "in-network" providers may not be the lowest cost providers for any given service. Dividing the accountability between the ordering and delivering provider as described above is more appropriate unless all three of the following conditions are met: (1) there are multiple providers who could deliver the service the physician wanted to order at an acceptable level of quality, (2) the physician knows the actual prices paid to those different providers (or the relative differences between the prices/payments), and (3) the physician has the ability to determine which provider is used to deliver the service.

It could also be argued that the physician or provider *delivering* the service should have some accountability over the number and type of services ordered, not just the relative cost of the services. However, this would depend on whether the delivering provider knew enough about the patient to determine the appropriateness of the order and whether the provider had the legal ability to modify an order from another physician.

### *Services Resulting from Referrals*

In many cases, a physician or other provider does not directly *order* a *specific* service from another provider, but *refers* the patient to the other provider asking either for advice about how the referring provider should treat the patient or asking the other provider to directly treat a particular health condition the patient is experiencing. Here again, one should assign a portion of the spending on these referrals to the referring provider in addition to the spending on services that the referring provider directly delivers or orders, otherwise there could be an incentive for a provider who wants a patient to receive an expensive service to refer the patient to another provider who would order and deliver that service.

In this case, the referring physician can be held accountable for making the referral that led to any services the second provider delivered or ordered, but it is not reasonable to hold the referring physician accountable for the *total* spending on those subsequent services. The physician or patient may or may not have had a choice about which provider to refer to, but even if they did have a choice, neither the physician nor the patient likely had any knowledge of how frequently the other provider ordered tests or performed procedures or how much that other provider is paid for their services. (One of the benefits of creating meaningful methods of assessing the cost and efficiency of different providers is to help physicians and patients make more informed choices about where to refer for services.)

For services resulting from referrals, the spending per patient can be disaggregated as follows:

- Since the physician making the referrals controls the number of referrals, the *expected spending per referral* would be assigned to the *ordering physician*. The "expected spending" per referral would be based on (1) the average types and numbers of services per patient that similar providers (either nationally or locally) delivered and ordered, and (2) the "reasonable" amount of spending on those services, as defined earlier.
- The *provider who received the referral* would then be accountable for the *deviation between the actual spending and the expected spending level* for all services that provider delivered directly in response to the referral. (If the provider who received the referral ordered services for the patient from a third provider, then, consistent with the methodology described earlier, the ordering provider would be accountable for the portion of the deviation based on what they *expected* the cost of the services from the third provider should be, and the third provider would be accountable for the portion of the deviation based on their actual payment vs. the expected payment.)

### *Which Physician Ordered the Service or Made the Referral?*

This allocation of accountability depends on being able to identify which physician or other provider ordered a service from another provider or made the referral to the provider who did order the service. In health insurance plans with HMO benefit designs where patients are required to both have a primary care physician and to have an order or referral from their primary care physician before seeing any specialist for non-emergency care, all initial services from specialists could be assumed to be orders or referrals from the patient's PCP. However, in traditional Medicare and in health insurance plans with PPO benefit designs, a patient is not required to have a PCP and even if they do, that PCP may not have any involvement in or knowledge about services the patient received from other physicians.

Fortunately, standard claims forms are already designed to capture information about the physicians who ordered or made referrals for services. Line 17 of the CMS-1500 Claim Form that is used by Medicare for physician billings includes a data field for the "Name of Referring Provider or Other Source." Moreover, the instructions in the Medicare Claims Manual state that a provider filing a claim for a service should "enter the name of the referring or ordering physician if the service or item was ordered or referred by a physician" in this data field, and the Claims Manual goes on to mandate that "All claims for Medicare covered services and items that are the result of a physician's order or referral shall include the ordering/referring physician's name."[94]

CMS is currently using this information in a manner similar to Spending Category 3 as part of the Supplemental Quality and Resource Use Reports it is providing to physicians. The reports include one table entitled "Breakdown of Episode Costs from Claims Billed, Ordered, or Referred by Eligible Professionals Within Your Medical Group Practice" and a separate table entitled "Breakdown of Episode Costs from Claims Billed or Ordered by Eligible Professionals or Facilities Outside Your Medical Group Practice."[95]

It is not known how accurate the information currently collected on ordering physicians is, but if it is *not* accurate, providers would have a much stronger incentive to ensure it *is* accurate if the information is used for develop-

ing spending measures. Using the methodology described above, a provider who delivered a service in response to an order from another physician would only be held accountable for any difference between their payment for the service and the payment amount determined to be reasonable based on national or local averages. However, if the provider delivered the service without documentation of who ordered it, they would be assigned the full cost of that service, so if they did not order the service, they would have an incentive to document who *did* order it. Ordering physicians would similarly have an incentive to ensure that the information on orders was accurate so they would not be held accountable for services they did not order.

Although the same data field appears on the National Uniform Claims Committee Form 1500[96] used by commercial health plans, many health insurance companies do not require that the field be completed other than by certain providers such as testing laboratories. Since most physicians participate in Medicare and complete this same data field for Medicare claims, it is not known whether they are also doing so on commercial claims even without it being enforced. Physicians would be less likely to resist the extra work in completing this field on all claims, and health plans should be less resistant to enforcing its completion, if the information begins to be used to produce more accurate and actionable spending analyses.

If the accuracy of the information about the ordering/referring physician is in doubt or if the information is not present at all, an alternative would be to include any service that is delivered by another provider within a defined period of time after the provider being measured delivered an evaluation and management (E&M) service that is included in Spending Category 1, if that other service has the same or a similar diagnosis code recorded as a diagnosis recorded for the E&M service. This is the same general approach that would be used by an episode grouper except that only services that *follow* an E&M service could be considered to have been ordered by the provider that delivered the E&M service.[97]

Where there is ambiguity, services could be including in spending tabulations for all physicians who *could* have ordered the service for a patient; although this would be more inaccurate than knowing who actually did order the service, it would be no more inaccurate than the way measures of total spending are currently being attributed to physicians, and it would create the incentive to record the information on who actually ordered services in order to generate more accurate measures in the future.

## Spending Category 4: Related Services

The fourth category involves spending on services that are clinically related to services that were delivered or ordered by the provider being measured, but are not included in any of the previous three categories. These are services that are:

- received by the *same patient;*
- occur *simultaneously with* or *after* the services that were delivered or ordered by the provider being measured;
- involve either a *similar diagnosis* to the diagnoses associated with the services the provider delivered or ordered or involve a *complication that could have resulted from* the services the provider delivered or ordered; and
- occur within *a timeframe* reasonably related to the services delivered by the provider being measured.

If a patient has formally designated a primary care physician or specialist to manage their care for one or more conditions or if the patient has formally been assigned to a physician through one of the processes discussed earlier in conjunction with Spending Category 1, then any services related to the conditions those physicians were managing should be included in Spending Category 4, even if the physicians did not explicitly order them.

A service would only be included in Spending Category 4 if the provider being measured did *not actually deliver the service*. The same service will also be included in Spending Categories 1-3 in a spending analysis for the provider who *did deliver* the service. As will be discussed further in Section IV-B, these two providers have shared accountability for this service, since one may have been able to prevent the need for the service and the other determined what service to deliver when the need occurred.

*EXAMPLE: A surgeon at an academic medical center performs knee surgery on a patient. After the patient is discharged, he develops an infection at the site of the incision, and his primary care physician admits him to the community hospital to treat the infection. The second hospital admission was not ordered by the surgeon, but it is likely related to the original surgery, so it would be included in Spending Category 4 for the surgeon as well as in Spending Categories 1-3 for the primary care physician and community hospital.*

*EXAMPLE: A primary care physician refers her patient to a cardiologist for evaluation of a potential heart problem. The cardiologist orders a stress test, and the test shows evidence of heart disease. Based on the test, the cardiologist sends the patient to an interventional cardiologist who does a cardiac catheterization and decides to use a stent. The stenting procedure was not ordered by the primary care physician, but it indirectly resulted from the referral the PCP made to the first cardiologist, so it should be included in Spending Category 4 for the PCP and in Spending Category 1 for the interventional cardiologist. (Ideally, the PCP would play a more active role in deciding which cardi-*

*ologist to refer patients to and in working with the cardiologist to help the patient decide what to do based on the stress test results.)*

This definition differs significantly from the way the methodologies described in Section II attribute total spending to providers. Under the most common attribution approaches, once the methodology identifies a primary care physician or other physician to whom a patient's spending should be attributed, *all* of that patient's spending is attributed to the physician, whether it is has any relation to the care the physician was providing or should have been providing or whether the physician had any ability to influence that spending. The variations in the attribution methodologies increase or reduce the probability that a particular patient's spending will be attributed to a physician, but once a patient is attributed to a physician, *all* of their spending is attributed. Under the approach defined in this section, only the spending that *chronologically follows* and is *clinically related to* services a provider delivered or ordered would be assigned to that provider.

Although spending analyses that attribute episodes of spending rather than total spending to providers also narrow the range of spending that is assigned to a provider, the above definition differs from the episode definitions used by typical episode groupers in three ways:

- Since the focus in the spending analysis is on a particular *provider*, only services that followed the *initial involvement with the patient by that provider* are included in Spending Category 4. In contrast, episodes are defined with a focus on the *patient*, so they will include services from some providers that precede the initial involvement of the provider who is being measured in the spending analysis. For example, if a patient experienced a complication after a hospitalization and had to be admitted to a different hospital to treat that complication, both hospitalizations would likely be included in the same episode, but the *original* hospitalization should not count as "related spending" for the hospital and physician that treated the complication.[98] Similarly, most "episodes" for chronic disease are defined in terms of services related to the chronic disease that are delivered during a twelve month period, but a hospitalization for a chronic disease exacerbation that precedes the initial involvement of a physician in treating the chronic disease should not be treated as "related spending" for that physician simply because it is included in the overall episode for the patient.
- In addition, since the focus is on providers, the same service can be included simultaneously in the spending assigned to two different providers (although it will be included in different categories of spending for each provider).[99] Episode groupers are generally designed to assign services to unique, mutually exclusive categories.
- For acute conditions, Spending Category 4 should include services that are related to the services delivered in Spending Category 1 even if they occur several months later. Episode groupers will generally assign services to two different episodes if the services are separated by more than a pre-defined length of time in which no related services occur (the "clean period"). As a result, an episode may fail to include delayed follow-up services or complications for which the same provider should appropriately be accountable.[100]

It is important to include truly related services in a measure of the total spending associated with a provider, because otherwise one provider may artificially appear to be "lower cost" than another provider when the second provider is actually spending more on a patient's initial care in order to prevent errors and complications, reduce hospital readmissions, etc. Figure 12 shows a simple example of how one hospital might appear to have lower per patient spending than another when looking only at the initial hospital stay, but when the cost of hospital readmissions are factored in, the total cost to the payer or patient is actually higher for the "lower cost" hospital. However, it is equally important to *exclude unrelated* services, otherwise a provider may appear to have high spending simply because their patients happen to be more likely to see other providers for different conditions or to see other providers who deliver or order many services.

In situations in which multiple providers are involved with the patient during the same time period, it may be impossible to definitively assign services in Category 4 to any particular provider. Rather than using the kinds of arbitrary rules described in Section II to attribute all of the spending to one provider and none to the others, it makes more sense to either (1) assign the spending to *all* of the involved providers, or (2) *apportion* the spending according to some measure of the relative involvement of the

## FIGURE 12
### Impact on Total Spending Per Hospital Admission of Differences in Readmission Rates

| | Payment for Initial Admission | Rate of Readmissions | Payment Per Readmission | Average Total Payment Per Patient |
|---|---|---|---|---|
| Hospital 1 | $15,000 | 15% | $20,000 | $18,000 |
| Hospital 2 | $14,000 | 25% | $19,000 | $18,750 |
| Difference | -7% | | | 4% |

different providers.[101]

*EXAMPLE: A patient has a heart attack, falls, and breaks a leg. The patient is admitted to the hospital, and during the hospital stay, an orthopedic surgeon repairs the fracture and a cardiologist places stents in two arteries. A week after the patient is discharged, he develops pneumonia and his primary care physician admits him to the hospital to treat the pneumonia. The surgeon, the cardiologist, and the hospital staff could all potentially have done more to prevent the pneumonia or treat it sooner, so the spending for treatment of the pneumonia should be included in Spending Category 4 for each of them (or it should be allocated among them based on some measure of their relative involvement in the patient's care).*

As with bundles of services, the ideal would be for providers to take accountability in advance for related services through a payment based on a full episode of care. This is discussed further in Section VI.

## Spending Category 5: Unrelated Services

For spending reports designed to help individual providers or groups of providers better manage the care of their patients, an optional final category would include the spending on everything else their patients received – services that had no logical, clinical connection to what the provider being measured did or could have done with respect to the patients.

In general, unless the provider has explicitly agreed to manage total spending for a pre-defined set of patients, the provider should not be held accountable for spending on the unrelated services in Spending Category 5. However, even if the provider has not taken accountability for total spending, it may be helpful for the provider to know that its patients have received these services and to make adjustments for the way care is being delivered in response. *Reporting* on Spending Category 5 allows a provider to see what is being spent on their patients without holding them directly *accountable* for it. Everything that would be included in Spending Category 5 for a particular provider would be included in Spending Categories 1-4 for other providers who would have the ability to influence it, so failing to hold the measured provider accountable for services in Spending Category 5 does *not* mean that *no* provider will be accountable for them.

*EXAMPLE: A patient who is severely injured as a passenger in an automobile accident will likely receive a large number of expensive services, but no primary care physician or specialist who was involved with the patient prior to the accident could have done anything to prevent it and the PCP would likely have little or no ability to control the costs of treatment for the injuries sustained in the accident. These services could be included in Spending Category 5 for the PCP for informational purposes, but not for accountability. The services would also be included in Spending Category 1 for the physicians and hospital that treated the patient for the injuries in the accident.*

## The Hidden Category: Services Without Associated Spending or Data

It is important to recognize that the above five categories only measure services that someone – the patient, Medicare, a health plan, an employer, etc. – *pays* for and for which *data are available*. As explained in Section III, if data are not available on certain services that can affect the need for or use of other services, erroneous conclusions could be drawn from analyses based on the more limited data that are available.

For example, as noted in Section III, access to data on pharmacy spending is likely very important for developing fair comparisons among many types of providers and for developing actionable information on ways to reduce overall spending. If data on pharmacy spending are only available for a subset of patients, it may be appropriate to carry out supplemental spending analyses limited to that subset of patients in order to determine how the results of spending analyses differ when more complete data are used.

In addition, many patients receive valuable services from a provider that the provider incurs costs to deliver but for which there is no payment and therefore no measured spending. For example, most patients likely place high value on the ability to speak to their physician on the telephone when they have a health problem, have a question or concern about their treatment plan, etc. However, Medicare and most health plans do not pay for such phone calls.

If one physician practice spends more time delivering unpaid services to patients than others do, that physician practice will likely be incurring additional costs for those services that are not included in the spending measure. That practice may also be delivering fewer billable services (e.g., the unbillable time physicians spend on the phone with patients would reduce the time they could spend in billable office visits with those or other patients), which would reduce the payer's spending and the practice's revenues at the same time that the practice's costs are increasing from delivering the unpaid services. If delivery of these unpaid services results in less need for other services for the patients that would have been paid for, e.g., fewer potentially avoidable emergency room visits, then the physician practice might score better on a measure of per patient *spending* than other providers, but that would not reflect the true difference in the total *cost* of the services delivered. Since delivery of these uncompensated services may not be feasible for other providers and it may not even be sustainable for the provider who is currently delivering them without a change in the payment system, the low spending measure would not be a realistic benchmark to use in evaluating other providers.

To the extent that data on these services are collected by providers (e.g., in electronic medical records), they could be compiled and used to supplement data on paid services derived from claims data to create a more complete picture of the services and costs incurred in patient care. If they are not routinely collected, research projects may need to be organized to collect them on a sample basis to determine whether they would produce different results than analyses based only on claims data.

## Differences in Analyses Using Spending Categories From Common Attribution Methods

This approach to categorizing spending would result in very different and much more actionable information than the current approaches to attribution described in Section II.

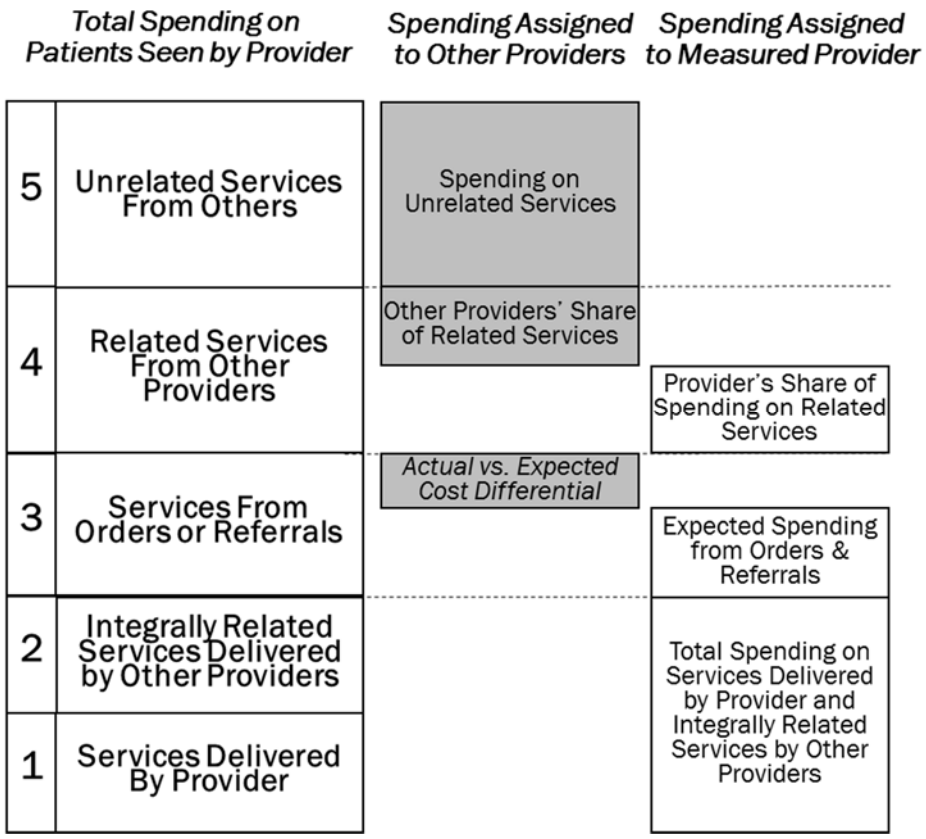
Figure 13 shows how the components of a report to a provider using this approach would differ from the reports typically being generated in conjunction with total cost of care measurement and accountability systems. In contrast to Figure 5 shown earlier, where there are more questions than answers about the involvement of the physician receiving the report and no clear indication of where or how spending could be reduced, a physician or other provider receiving the information shown in Figure 13 would know:

- That the spending in the first three categories resulted from services delivered directly by the provider, from services of other providers that are integrally related to those services, from services ordered by that provider, or from services that resulted from referrals the provider made; and
- The extent to which spending on ordered services was higher because the price of those services was higher than what is available from other providers.

Consider a hypothetical patient who receives the following sequence of healthcare services during the course of a year:

- In January, the patient visits his primary care physician (PCP) complaining of mild chest pain while exercising. The primary care physician orders a cardiac stress test to help determine if the patient is at risk of a heart attack.
- In February, the stress test is performed and the cardiologist who reviews the results determines that there is no indication of significant coronary artery blockage. The cardiologist sees the patient in the cardiologist's office to explain the results, determines that the patient has risk factors for a future heart attack, and orders medication to try and reduce that risk.
- The patient has also been having lower back pain. The patient does not consult with the PCP about this problem, but contacts a neurosurgeon directly and schedules an appointment in March. The neurosurgeon evaluates the patient and recommends spinal surgery.
- The patient receives spinal surgery at the hospital in April. An anesthesiologist provides the sedation during surgery.

### FIGURE 13
### Breaking Down Total Spending on Patients Into Five Spending Categories



- After discharge, the patient decides to go to a skilled nursing facility for inpatient physical therapy and rehabilitation rather than for outpatient physical therapy, and the patient's insurance approves that service.
- When the patient was being discharged from the hospital, the neurosurgeon recommended that the patient see a primary care physician regularly after completion of therapy. The patient decides to use a different primary care physician than the one he had seen in January.
- The patient visits the new primary care physician in June and again in October, and the new PCP makes sure that the patient is up to date on all preventive care. The PCP finds that the patient has not had recommended screening for colon cancer and orders a colonoscopy.
- A gastroenterologist performs the colonoscopy in November and finds no polyps or evidence of cancer. The gastroenterologist performs the colonoscopy in a hospital and uses an anesthesiologist to administer sedation.

This patient has received services from a total of seven different physicians. As shown in Figure 14, accountability systems that attribute spending based on primary care visits such as those described in Section II would attribute all of the services from all of those physicians to the patient's new primary care physician, including the stress test and the back surgery, even though those services occurred before the new primary care physician first met the patient. None of the services would be assigned to the other physicians who actually delivered or ordered them (the two anesthesiologists are not shown separately). This is because current attribution methodologies assign *all spending* for the *full year* to the *primary care physician* who had the *majority* (or plurality) of *primary care office visits during the year.* Since the patient had more visits with the new PCP than with the PCP he saw in January, all of the spending on the patient is attributed to the new PCP.

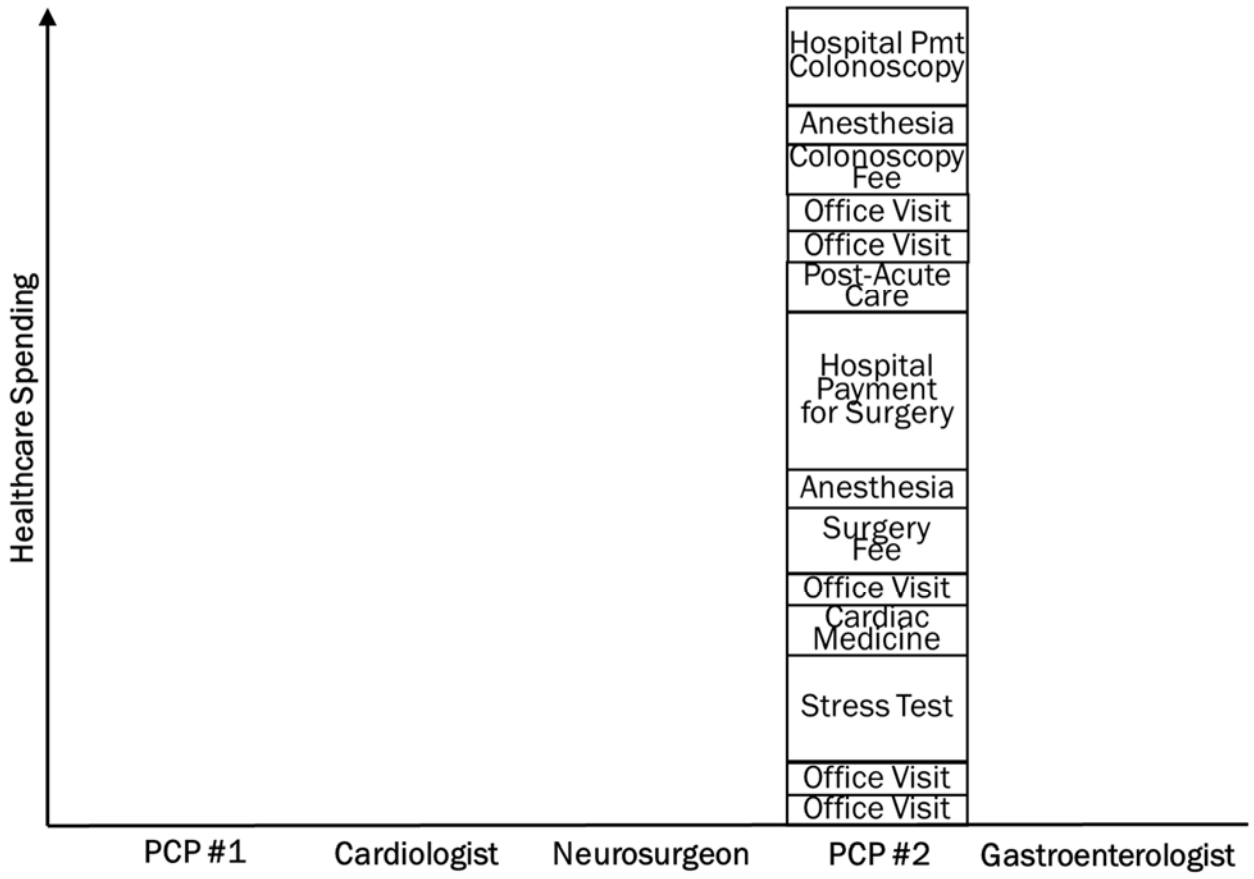
If the attribution were based on which physician of any specialty had the majority of office visits with the patient, the patient might not be attributed to anyone at all, since the patient had a total of five office visits with physicians during the year, and none of the physicians had a majority of those visits. If the attribution were based on the physician who billed the most expensive services, all of the spending, including the cardiac testing and the colonoscopy, would likely be attributed to the neurosurgeon, even though the cardiology care occurred before the neurosurgeon was involved and the neurosurgeon likely did not see her role as managing the patient's preventive care for colon cancer.

As shown in Figure 15, the spending category system described in this section would assign individual components of spending to each of the physicians who had control over each service, rather than assigning total spending to a single physician or to no physician at all:

- The visit in January to the patient's original primary care physician would be included in Spending Category 1 for that PCP. A portion of the cost of the cardiac testing the PCP ordered would be included in Spending Category 3 for that PCP.
- The remaining portion of the cardiac testing would be included in Spending Category 3 for the cardiologist. The office visit with the cardiologist would be included in Spending Category 1 for the cardiologist, and the full cost of the heart medications ordered for the patient would be included in Category 3 for the cardiologist.
- The neurosurgeon's fees would be included in Spending Category 1 for the neurosurgeon. The spending for the anesthesiologist and the hospitalization would be included in Spending Category 2 for the neurosurgeon. The cost of the post-acute care would be included in Spending Category 4 for the neurosurgeon, since it was related to the procedure the surgeon performed in the hospital, but the surgeon did not directly order inpatient rehabilitation or choose which facility was used.
- The two office visits with the second PCP are included in Spending Category 1 for that PCP. A portion of the cost of the colonoscopy ordered by the PCP is included in Spending Category 3 for the PCP.
- The remainder of the cost of the colonoscopy, including the cost of the anesthesiologist and the hospital, is included in Spending Category 3 for the gastroenterologist, who is responsible for the fact that the colonoscopy is performed in a hospital using an anesthesiologist, making it more expensive than if it were done in the gastroenterologist's office using an alternative form of sedation.

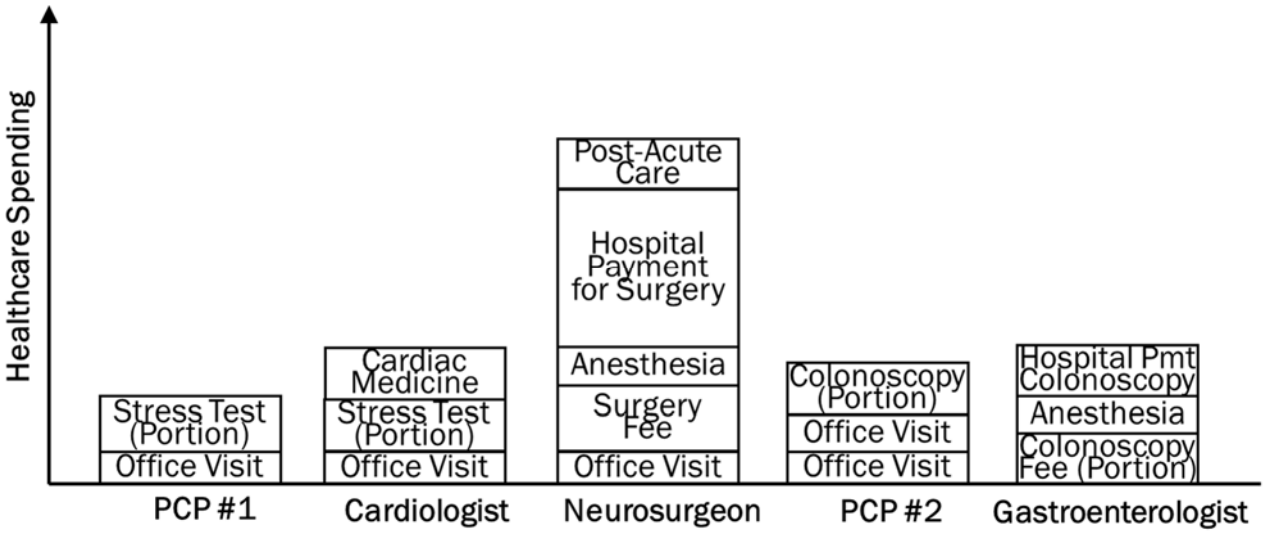
## FIGURE 14

Attributing Total Spending to the Primary Care Physician With the Plurality of Visits




## FIGURE 15

Assigning Components of Spending to the Providers Best Able to Control Them



## B. Identifying Subsets of Services Where Spending Can Be Reduced

The breakdown of spending described in the previous section better identifies which providers have the ability to control or influence different aspects of spending, but it gives only a limited indication as to which aspects of spending could be *reduced*. To better identify opportunities to reduce spending without harming patients, it is desirable to further disaggregate Spending Categories 1-4 into four subcategories:

(a) Services required to meet quality standards

(b) Services that are potentially avoidable

(c) Services needed to address potentially preventable conditions

(d) All other services ("typical services")

### Subcategory (a): Services Required to Meet Quality Standards

Before identifying services that can be reduced, it is important to identify those services that should *not* be reduced and may well need to be *increased* if they are currently being underused.

CMS and most commercial payers have established pay-for-performance programs for physicians, hospitals, and other healthcare providers that reward or penalize them based on the rate at which they deliver or order various services that evidence has shown are desirable or necessary for good patient care. For example, many pay-for-performance programs use the Colorectal Cancer Screening Measure (NQF #0034) to reward physicians whose patients have high rates of receiving screening colonoscopy, and they use the Comprehensive Diabetes Care: LDL-C Screening Measure (NQF #0063) to reward providers whose diabetic patients are tested for LDL cholesterol annually.

All else being equal, however, higher rates of blood tests and colonoscopies will increase spending, particularly in the short run. It would be inappropriate to penalize a physician or other provider for being "inefficient" because of higher spending on the same services the provider is being encouraged to provide or order based on quality standards.[102] Conversely, one would not want to label a physician as "efficient" because they failed to deliver or order recommended services for their patients. Simply creating composite measures that include both quality and spending metrics can inappropriately force providers to compare the rewards or penalties associated with quality measures vs. spending measures to determine whether it is better to deliver more recommended services (thereby generating rewards on quality measures but penalties on spending measures) or to reduce spending by delivering fewer recommended services.

The solution is to separately tabulate spending on the services associated with the quality measures within each of the Spending Categories defined in the previous section, so that spending on services that providers are specifically being encouraged to deliver is measured separately from other services. Since most of the quality measures currently in use are process-based measures computed using claims data, the claims that are used to calculate the number of patients in the numerators of these measures (i.e., the number of patients who received a recommended service) can also be used to identify the *spending* on the recommended services. For example, if a blood test ordered by a primary care physician qualified the patient to be included in the numerator of the Comprehensive Diabetes Care: LDL-C Screening Measure, the spending on that test would be included in subcategory (a) of Spending Category 3 for that PCP.

There are inherent limitations in using claims data to determine whether a particular service is desirable for a particular patient. However, these limitations create the same degree of inaccuracy in the quality measures that use the claims-based specifications as they would in defining measures of spending associated with recommended services. Efforts are being made nationally to calculate quality measures using information from electronic health records rather than claims data, and as this progresses, new methods will be needed to identify the spending associated with evidence-based services.

Ideally, quality should be measured based on patient outcomes, not on services. As quality measurement moves away from process-based measures, it will be necessary to compare performance on spending and outcomes jointly, rather than separately. This is discussed in more detail in Section E below.

The fact that a service is recommended by quality standards does not mean that it cannot be delivered in a lower-cost way. The methodologies described in Section IV-A for breaking down spending in Categories 1-3 based on utilization and price can be applied separately to this subcategory in order to determine whether colonoscopies, blood tests, etc. are being delivered as efficiently and affordably as possible.

### Subcategory (b): Services That Are Potentially Avoidable

A second subcategory of services can be described as "potentially avoidable." These are services such as MRIs for lower back pain, cardiac stress tests, and Cesarean sections that may provide significant benefit to some patients but relatively little benefit to others and in some cases may result in harms to the patient that outweigh the benefits.

As part of the Choosing Wisely® campaign coordinated by the ABIM Foundation, over fifty medical specialty societies have each identified five frequently used tests, procedures, and other services that may not be necessary for individual patients.[103] 3M Health Information Systems has developed a methodology for identifying "Potentially Preventable Services;" these are tests, procedures, and medications ordered by primary care physicians or specialists in ambulatory care settings that may not provide useful information for diagnosis and treatment.[104]

It is important to note the use of the word "potentially." It is generally not possible to determine that a service was *definitely* avoidable, particularly with the limited information available in a claims data system.[105] Moreover, depending on the types of patients a physician or other

provider is caring for, a higher rate of these services may well be justified, so the mere fact that spending is higher for one provider than another does not justify rewarding or penalizing the provider based on that. (Addressing differences in patient populations is discussed in Section IV-C below.) However, because these services are known to be subject to overuse, separately tabulating spending on them can help providers and payers identify potential opportunities to reduce spending without harming patients.[106]

## Subcategory (c): Services to Treat Potentially Preventable Conditions

The third subcategory involves services that are also potentially avoidable, but avoidable in a different way. Some patients receive services to treat health conditions where the *health condition itself* could potentially have been prevented if additional or different services had been delivered *at an earlier point in time*. For example, the services needed to treat an infection at a surgical site might have been avoided if better infection control procedures had been used when the surgery was performed, and a hospital admission for an asthma attack might have been avoided if the patient had received additional services from their PCP or a specialist to help them successfully manage their asthma. Research has shown that these are the types of spending that patients are most likely to be concerned about, since they are associated with health problems that could have been prevented had the patient received better quality care.[107]

As with the potentially avoidable services subcategory, these conditions are only "potentially" avoidable, because it will generally not be possible to determine with certainty that a change in care at an earlier point in time would definitely have prevented the condition.

### *Defining Potentially Preventable Conditions*

Several systems have been developed to define services associated with potentially preventable conditions. As part of its work to develop PROMETHEUS episode payments, the Health Care Incentives Improvement Institute has developed definitions of "potentially avoidable complications" associated with particular procedures and hospitalizations and with chronic disease management.[108] Several measures based on these definitions of potentially avoidable complications have been endorsed by the National Quality Forum.[109]

Similarly, 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."[110] 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.

### *Assigning the Spending on Potentially Preventable Conditions to Providers*

For potentially preventable conditions, the provider who delivered the service to *treat* the condition may or may not have been the provider who could have *prevented* the condition from occurring. For example, if a patient gets an infection following a procedure performed in a hospital, the patient may be treated for the infection in a different hospital and by a different set of physicians than the physicians involved with the original procedure.[111] If a patient who is being treated by a primary care physician for a chronic disease experiences a serious exacerbation and has to be admitted to a hospital, the primary care physician may not be involved with the patient's care in the hospital.

Consequently, a service in Category 1 should only be categorized in subcategory (c) if the provider *delivering* the service *was involved with the care of the patient at a previous point in time and in a way that could have caused or prevented the condition*. If the provider was not involved in that way, the service the provider delivered should still be included in Category 1, but it should be placed in one of the other subcategories. If the service was required/recommended for the patient's condition, then the service should be placed in subcategory (a); if the service was known to be frequently overused, it should be placed in subcategory (b); otherwise, it should be placed in subcategory (d). Conversely, if a provider *was* involved with the care of the patient in a way that could have *prevented* the condition, but that provider did not deliver any services to *treat* the condition, then those services should be included in Category 4 ("related services") for that provider, and within Category 4, the services should be grouped in subcategory (c) for "services associated with potentially preventable conditions."

*EXAMPLE: If a patient had surgery at one hospital, developed an infection after discharge, and was admitted to a different hospital and treated for the infection by physicians who were not involved in the surgery, then the hospitalization and physician services to treat the infection would not be classified as "services associated with potentially preventable conditions" for the hospital and physicians who treated the infection. The infection was potentially preventable, but not by those who treated it, and once the infection occurred, it needed to be treated. Consequently, the spending should be included in subcategory (d) of Spending Category 1 for the hospital and physician who treated the infection, but it should also be included in subcategory (c) of Spending Category 4 for the surgeon and the hospital where the surgery was performed.*

*EXAMPLE: If a patient with a chronic disease is receiving care from a primary care physician and has to be hospitalized for an exacerbation of the disease, the spending on that hospitalization should be included in Category 4(c) for the primary care physician even though the physician did not admit the patient to the hospital or treat the patient during the hospitalization.*

In some cases it will be reasonably clear that a problem experienced by a patient is a complication resulting from a specific healthcare service they received previously, but in other cases it will not be clear. For example, if a patient develops an infection in their hip after hip surgery, it is highly likely that the complication was related to the hip surgery. However, if a patient is diagnosed with pneumonia after being discharged following a brief hospital stay, it may be difficult to determine whether the patient contracted the pneumonia *during* the hospital stay, *after* discharge, or perhaps even *before being admitted* to the hospital. Moreover, whenever multiple providers have been involved in a patient's care, it may also be difficult to directly link a complication with a particular service or provider. For example, if a patient who is treated in the hospital and discharged to a skilled nursing facility for rehabilitation has to be readmitted to the hospital, it may be impossible to determine whether the readmission was due to something that occurred in the hospital or in the nursing facility or both (or due to something that *should have* occurred in either facility but did *not* occur). In these cases, the services for the avoidable condition can be assigned to all of the providers who were involved or apportioned among them in some way.[112]

In many cases, a potentially preventable condition does not occur because of complications of a service that *was* provided, but from complications resulting from the *lack* of a service that *could or should have been provided*. For example, many patients with chronic diseases are seen in an emergency room for exacerbations of their chronic disease; these exacerbations generally require emergency treatment, but the exacerbations themselves could have been prevented in many cases by better care from a primary care physician or a physician specializing in that chronic disease. However, if the patient has not formally designated a primary care physician or specialist to help them manage their chronic condition, it may not be clear which, if any, physician should be held accountable for the preventable visits to the emergency room and any hospitalizations which follow. In these cases, the service could be assigned to the provider(s) who (1) most recently provided services to the patient and (2) could have provided services that potentially could have prevented the condition, if there are any such providers. For example, if a primary care physician treated a patient for their chronic disease within a certain number of months prior to a hospitalization related to the chronic disease, that avoidable hospitalization could be assigned to that primary care physician. If both a primary care physician and a specialist in the patient's disease saw the patient prior to the potentially avoidable event, then the services associated with the avoidable event could be assigned to both of the physicians or apportioned between them based on their relative involvement in the patient's care.

In cases where the preventable condition cannot be reasonably associated with any provider who could have caused or prevented it, then the spending *should not* be assigned to any provider other than the provider who delivered the treatment. In these cases, the only other "accountable" entity is the patient and/or the patient's health insurer (including Medicare or Medicaid), since one or both of them has not arranged for the patient to receive the kind of care that would have prevented the undesirable condition.

### Subcategory (d): All Other Services ("Typical Services")

Any service that does not fit into the other three subcategories would fall into subcategory (d). For these "typical" services, there may not be enough evidence about appropriateness to classify them as either recommended services (and thereby appropriate for subcategory (a)) or potentially avoidable services (and thereby appropriate for subcategory (b)), and there may be nothing to suggest that the condition involved could have been prevented by the provider (thereby making the services appropriate for subcategory(c)). Even so, variation among providers in the types and number of "typical" services they use for similar patients could indicate opportunities for savings and areas where research to determine appropriate use criteria are needed.

### How Subcategories Would Produce More Actionable Information

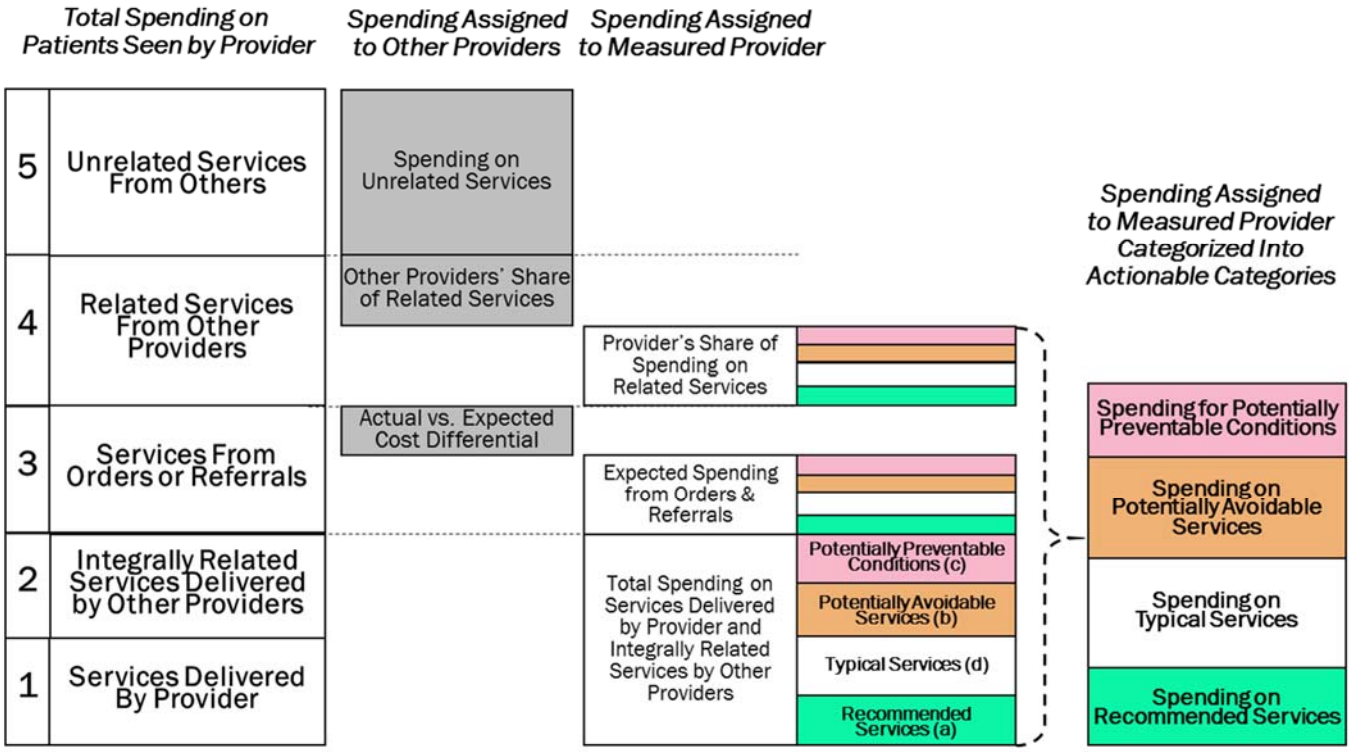
Figure 16 shows how the disaggregated categories would enhance the breakdown shown at the end of Section IV-A by indicating how much of the spending assigned to the provider was recommended for the patient, how much was potentially avoidable, and how much was associated with conditions that could potentially have been prevented.

Consider a hypothetical patient with COPD (emphysema) who receives the following sequence of healthcare services during the course of a year:

- In January, the patient visits her primary care physician (PCP). The PCP updates the patient's prescriptions for bronchodilators and asks the patient to schedule another visit in July.
- In February, the patient visits the PCP complaining of lower back pain. The PCP orders an MRI, which shows no spinal problems. The PCP tells the patient it is likely just a minor back strain and recommends an over-the-counter pain reliever.
- In April, the patient contracts a cold and has increasing difficulty breathing. She does not call her PCP, and when her breathing difficulty becomes extreme, she goes to the emergency room and is admitted to the hospital.
- The patient is treated for the COPD exacerbation by a hospitalist and pulmonologist at the hospital and discharged after five days.
- In May, just a few weeks after discharge, the patient experiences another exacerbation and she is hospitalized again. A different hospitalist treats her in the hospital and arranges for the patient to see a new primary care physician after discharge.
- The patient visits the new PCP in June, who adjusts the patient's medication and sends a nurse to visit her in her home to make sure she is using the medication properly. The PCP schedules the patient for a return visit in September, but urges the patient to call

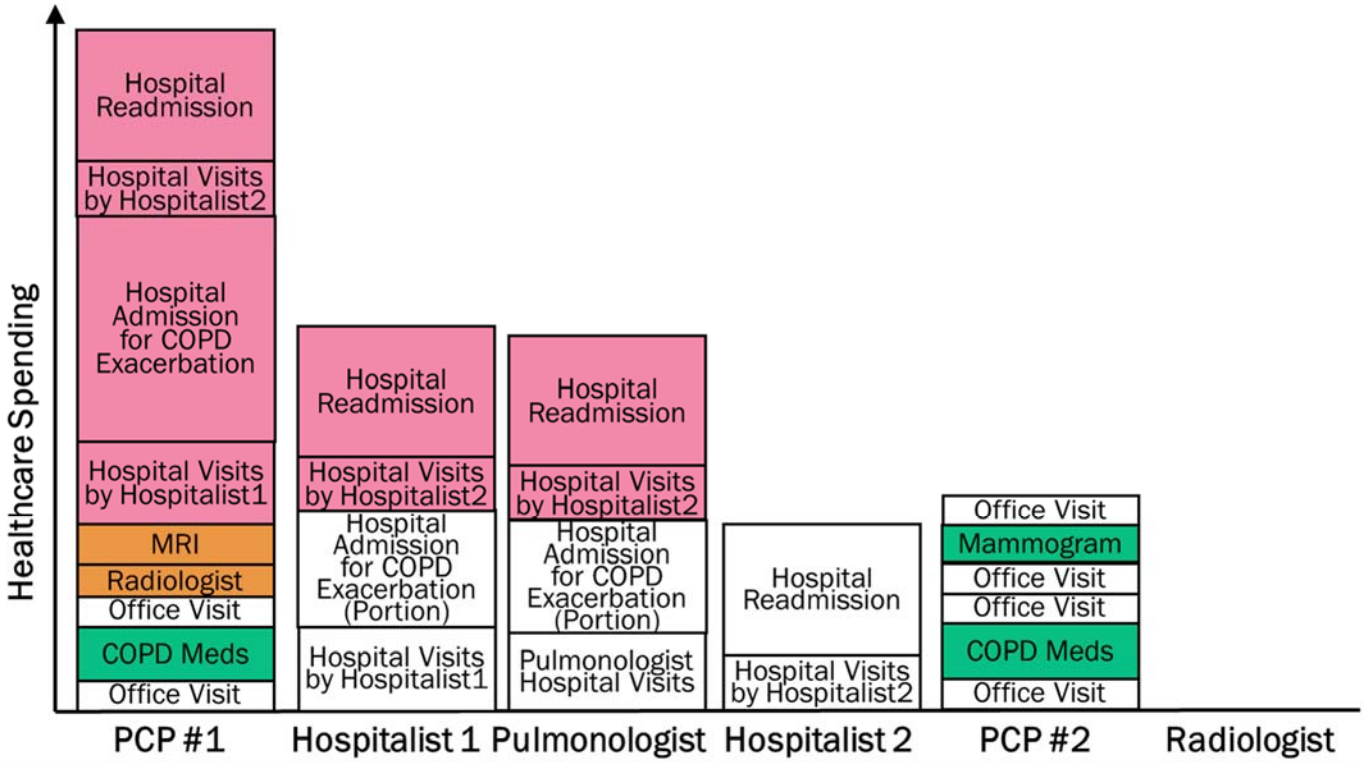
## FIGURE 16

Disaggregating Total Patient Spending Into Actionable Categories for Each Provider




## FIGURE 17

Assigning Opportunities to Reduce Spending to the Providers Best Able to Achieve Them



immediately at the first sign of a cold or any problem in breathing.

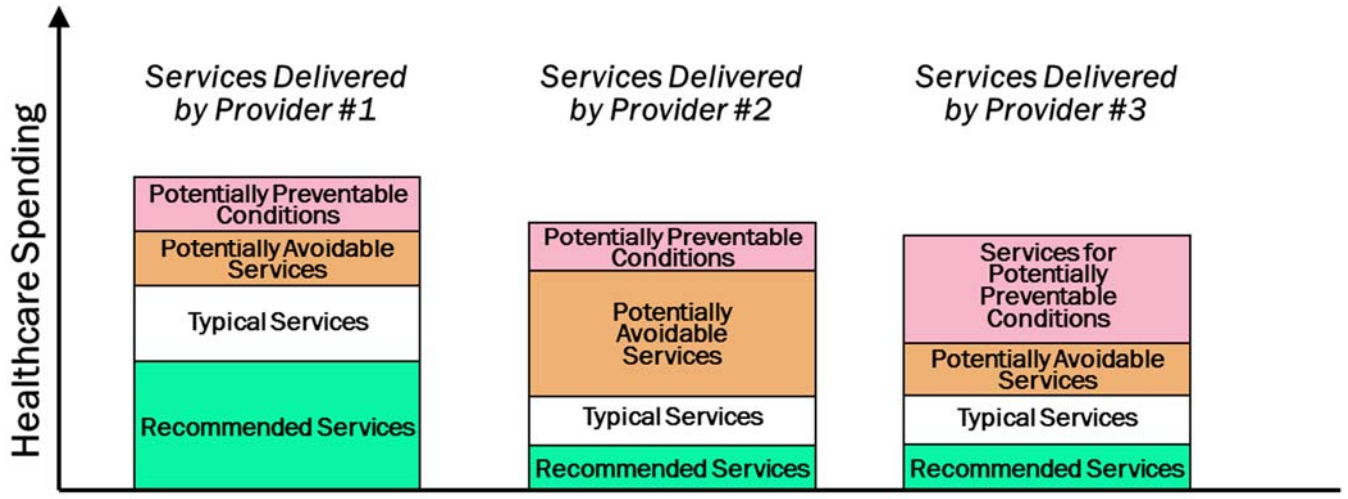
- At the September visit, the PCP determines that the patient is managing her COPD well, but the PCP notices that the patient has not had a mammogram recently. The PCP orders a mammogram for her.
- The patient obtains the mammogram in October and a radiologist determines that it shows no sign of breast cancer.
- The patient visits her PCP again in December and is again found to be managing her COPD well.

The patient has received services from six different physicians, none of whom were in a position to coordinate or control all of the services the patient received during the entire year. Nonetheless, typical accountability systems that attribute spending based on primary care visits would attribute all of the above services and spending to the patient's new primary care physician, including the two hospitalizations that occurred before the new primary care physician first met the patient, because the PCP delivered the majority of primary care services during the entire year.

As shown in Figure 17, the categories and subcategories described in Sections IV-A and IV-B would assign individual components of spending to each of the physicians who had the most control over them, and it would also break down the services into subcategories to show which services were potentially avoidable.

- The two office visits with the first PCP would be included in Spending Category 1(d) for that PCP and the three office visits with the second PCP would be included in Spending Category 1(d) for that PCP.
- The bronchodilators prescribed by each PCP would be included in Spending Category 3(a) for each of them, since they are recommended medications for COPD.
- The MRI ordered by the first PCP would be included in Spending Category 3(b) for that PCP, since MRIs for lower back pain are generally unnecessary. (This categorization does not imply that the MRI was *inappropriate* for this particular patient, since the PCP may have had reasons to be concerned about other potential causes of the pain.)
- The mammogram ordered by the second PCP would be included in Spending Category 3(a) for that PCP if it met current guidelines for appropriate breast cancer screening.
- The hospitalization (both the hospital care and physician services) in April would be included in Spending Category 4(c) for the first PCP (since that PCP could potentially have prevented the exacerbation that resulted in the hospitalization), and would also be included in Spending Categories 1(d) and 2(d) for the hospital, the hospitalist, and the pulmonologist (since the patient needed hospital treatment for the exacerbation once it occurred).
- The cost of the hospitalization and the second hospitalist's services in May would be included in Spending Category 4(c) for the first PCP, the first hospitalist, and the pulmonologist, since any or all of them could have provided services such as better patient education and transitional support following the first hospitalization that potentially could have prevented the second hospitalization. The hospital cost for the May hospitalization would be included in Spending Category 1(c) for the hospital since it also could have taken steps during the first hospitalization to prevent the condition that led to the hospital care it delivered during the second admission. The second hospitalist's services in May would be included in Spending Category 1(d) for the hospitalist, since that hospitalist had not previously been involved with the patient's care.
- No spending would be assigned to the radiologist, since the MRI and mammogram were ordered by the two PCPs and the prices of the MRI and mammogram and the radiologist's fees for interpretation are reasonable for those services.

FIGURE 18
Comparing Spending in Actionable Categories



The value of this kind of breakdown is further demonstrated by Figure 18, which shows the per-patient spending associated with the services delivered by three hypothetical providers. Provider #1 has the highest average spending, but its higher spending is due to greater use of services required for quality standards, not due to potentially avoidable services or services associated with potentially preventable conditions. In the cases of both Provider #2 and Provider #3, higher spending is due to either use of potentially avoidable services or treatment of conditions that could potentially have been prevented, and there may be underuse of recommended services.

## C. Addressing Differences in Patient Needs

The categories and subcategories of spending defined in the previous sections will help a provider find opportunities for reducing spending even if the spending information is not compared to that of other providers. However, since many services will be categorized as "typical" and even the potentially avoidable and preventable services are merely "potential" opportunities, additional insight into opportunities for reducing spending can be developed by comparing spending among providers to identify types of services that may be unnecessary as well as high-value services that are being underused. Moreover, if payers want to reward providers who spend less to treat their patients, they need a way to identify which providers have lower spending for the right reasons.

These comparisons will only be meaningful if they control for differences in the patients for whom different providers are providing care. A physician or other provider may have higher spending simply because their patients have more needs or needs that require more services or more expensive services, and if adjustments are not made for this in spending comparisons and payment systems, it could discourage physicians and other providers from taking on higher-need patients and make it more difficult for these patients to obtain the care they need. Moreover, the methods used to adjust for differences in patient needs must address the many weaknesses of current risk adjustment systems described in Section III-E.

### Disaggregating Spending into Subgroups of Patients with Similar Health Conditions

It would certainly be desirable if fair comparisons could be made simply by dividing spending by a single number representing the magnitude of a patient's needs, as risk adjustment systems based on risk scores do. However, as the discussion in Section III-E makes clear, patient needs are multi-dimensional and cannot be adequately captured by a single number. This is particularly true when the subsets of spending associated with different physicians and other providers are examined separately. As shown in the examples in Section III-E, two patients can have the same risk score as a result of very different combinations of health problems, but those different combinations of health problems will have very different impacts on the types of services providers in different specialties will need to deliver. In particular, some of the health problems will be appropriate for management by a primary care practice and some will not, and risk adjustment based on a single score cannot make that distinction.

One solution to this is to develop a different risk scoring system for each different health condition or each different episode. Under this approach, for example, if one were trying to compare spending among cardiologists, each patient would be assigned a cardiovascular risk score based on their presumed need for cardiovascular services, and then the spending on services delivered or ordered by each cardiologist would be divided by the average cardiovascular risk score for their patients in order to compute the risk-adjusted spending for cardiologists. The patient would be assigned a different risk score to measure their need for orthopedic services, cancer care, etc., and then spending by orthopedic surgeons, oncologists, and other physicians would be divided by the average orthopedic risk scores, oncologic risk scores, etc.

The advantage of this approach is that the patient characteristics that most affect needs within each specialty, condition, or episode can be identified and weighted appropriately, rather than using the same factors and weights from a more generic risk adjustment system. However, the disadvantages of this approach are:

- it assumes that for patients with the same health condition or episode, a scale can be defined such that two individuals with the same risk score will require services involving similar amounts of spending and that an individual with a higher (or lower) score will require services that cost proportionally more (or less). The most common approach to producing risk scores – linear regression analysis – assumes that patient characteristics affect the services they need in fairly simple linear ways, even though it is likely that the real interactions are more complex and non-linear.
- it leads to a potentially complex multiplicity of scoring systems. Since the same patient characteristics will affect different services in different ways, each characteristic must be weighted differently and combined with different sets of other characteristics in different ways depending on which type of specialty, condition, or episode is being examined. Individual patients could have dozens of different risk scores applicable to different conditions, different episodes, etc.

CMS is using this type of approach in its Hospital Value Based Purchasing Program and Physician Value Based Payment Modifier program. For each episode of care that is measured in these programs, a separate regression analysis is performed to select the patient factors and weightings that will be used to estimate which of the patients experiencing a particular type of episode would have higher spending within that episode.[113] Then the spending amounts for the episodes assigned to a particular provider are adjusted by applying the individual regression formulas to the characteristics of the patients involved with each episode.

An alternative and preferable approach is to compare spending separately for different subgroups of patients, with each subgroup defined such that patients in that subgroup would be expected to need similar levels of ser-

vices. The spending on the services that the patients in each subgroup received from a provider would then be compared to the spending on the services that patients in a similarly defined subgroup received from other providers.

There are several important advantages to this approach:

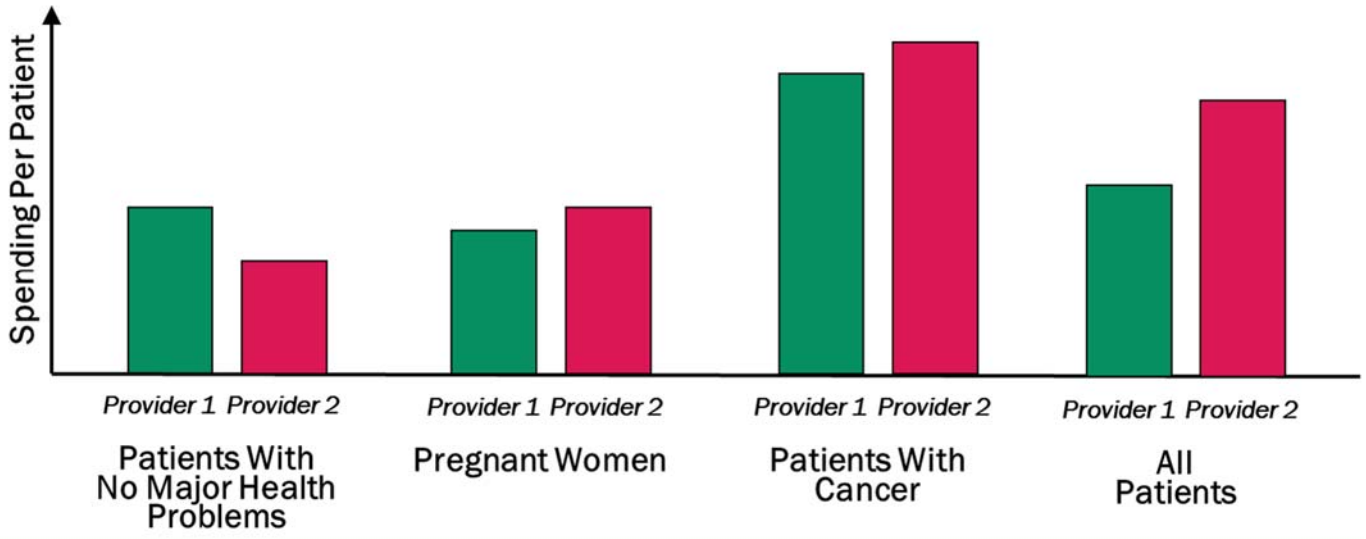
- it does not require there be any pre-defined relationship between the needs and spending for patients who fall into particular subgroups, as is required for regression-based risk scores.
- it is far easier for providers to understand the differences in the needs of patients in clinically defined subgroups (e.g., "women who are over 65 with diabetes and COPD" vs. "women who are over 65 with no chronic diseases") than if they are in subgroups defined by a formula-driven risk score (e.g., "patients who score between 2.3 and 2.5 using a weighted sum of factors measuring their age, gender, and number of chronic diseases" vs. "patients who score between 1.0 and 1.2 using that same formula").
- it allows spending for individual providers to be compared only for those patients for whom they would be expected to manage care. For example, patients with cancer who are undergoing chemotherapy treatment could be excluded from spending comparisons for primary care physicians.
- it enables identification of specific subgroups of patients where spending is high (or low), so that further analysis can be done to determine the factors causing that and so that interventions can be more effectively targeted on those patients where spending is high. By defining the subgroups in clinical terms, the physicians in the specialties relevant to the patients' conditions can work together to redesign care for those subgroups, whereas a patient group defined by having the same risk score could span almost every specialty. Moreover, the definitions of the subgroups will only change when there is a reason to believe that the relationships between patient characteristics and appropriate healthcare services have changed, rather than changing because a different set of variables and weights achieved a better predictive score in a regression analysis.

For example, as shown in Figure 19, if spending for patients with different types of health conditions were analyzed separately, it could reveal that a provider with lower risk-adjusted *total* spending had higher spending for *some* subgroups of patients. Total risk-adjusted spending per patient is higher in the second primary care practice than in the first because of higher spending for treatment of patients with cancer and higher spending for labor and delivery for pregnant women, even though those differences were caused by the way oncologists or obstetricians practice in that community not by the primary care practice, and even though the second primary care practice actually had lower spending for the types of patients where primary care could make an impact.

CMS has been using this approach for over twenty years in the way it pays hospitals for inpatient care through the Inpatient Prospective Payment System. Separate Diagnosis Related Groups (DRGs) are defined for groups of patients that are viewed to be similar in terms of the clinical characteristics that would require different types and intensities of hospital services. In the current MS-DRG system, there are two or three different DRGs defined for any major condition or procedure, based on the number and severity of health problems the patient has that would affect services and spending for treatment of that specific condition or delivery of that specific procedure.[114] CMS has indicated that having clinically meaningful groupings has been essential for helping providers manage costs effectively without harming patients.[115]

FIGURE 19

Comparing Spending Differences for Patients With Similar Characteristics



Fortunately, some risk adjustment systems have methods for grouping patients into clinically similar subgroups that can be used in this way for all types of spending, not just for inpatient hospital care. For example,

- The Adjusted Clinical Groups (ACG) system developed by Johns Hopkins University uses information on the duration, severity, diagnostic certainty, and origin of a patient's diagnoses to categorize each of the patient's health problems into one of 32 diagnosis clusters. Then, based on the patient's age, sex, and the diagnosis clusters applicable to them, the patient is assigned to one of 93 different ACG categories.[116]
- The Clinical Risk Group (CRG) system developed by 3M Health Information Systems uses a patient's diagnoses and past medical interventions to determine whether the patient has chronic conditions or acute conditions or both and the severity of those conditions.[117] This information is then used to assign the patient to one of 269 different "Base CRGs" and to one of up to 6 different severity levels, for a total of 1,080 potential different groupings.[118]

A useful feature of these systems is that they define multiple levels of aggregation and disaggregation, so that different types of physicians and provider organizations can look at the data in a way that is most relevant to them. Services and spending on patients in different subcategories can then be analyzed separately to determine where overuse or underuse is occurring and how to address it.

*EXAMPLE: The 3M Clinical Risk Group (CRG) system can group patients into 9 different "Core Health Status" categories which would help primary care physicians or multi-specialty physician groups better identify which subsets of patients are receiving unnecessary services or experiencing avoidable complications[119]:*

- *Healthy (i.e., no chronic or significant acute conditions)*
- *A significant acute condition*
- *A single minor chronic condition*
- *Multiple minor chronic conditions*
- *A single dominant or moderate chronic condition*
- *Multiple significant chronic conditions*
- *Three or more dominant chronic conditions*
- *A dominant or metastatic malignancy*
- *A catastrophic condition*

*The CRG system then also allows further disaggregation of patients in each Core Health Status category into more specific sub-categories which will be more relevant to determining which patients and which aspects of care represent opportunities for improving quality or reducing costs. For example, patients in the "dominant chronic conditions" categories can be subdivided into different "episode diagnostic categories" to distinguish whether the patients have congestive heart failure, COPD, or diabetes, since different conditions will be treated by different sets of physicians.[120]*

For providers with small numbers of patients, different levels of disaggregation also enable spending comparisons to be done only where there are statistically valid numbers of clinically similar patients. For example, if an individual physician has a diverse mix of patients across a large number of different categories, comparing the spending on a subset of the physician's patients *in each risk category* to spending by other providers on patients in the *same risk category* will be more valid than comparing spending on *all* of the physician's patients to that of other physicians.

### *Differences Between Clinical Categorical Models and Regression Models*

Not every risk adjustment system has the kind of underlying clinical logic that allows this type of disaggregated analysis. As noted in Section III, risk adjustment systems developed with linear regression methods choose the patient characteristics that are statistically best at predicting spending in the year in which the regression analysis was performed, even if there is no clinical logic indicating that patients who are different in terms of those characteristics should need different types or numbers of services. The regression analysis cannot distinguish between factors that have a causal relationship to services and those that have a spurious correlation. Moreover, since the regression parameters are estimated based on the overall population of patients, they may be far less successful at predicting services or spending for individual subsets of patients.[121]

It is certainly more complex to make multiple comparisons of spending for different categories of patients than to make one comparison by dividing total spending for all patients by a single risk score, but it is worth the effort because there is no way to know whether any single risk score is "right" or not. In fact, none of the major risk adjustment systems is very effective at doing what they are designed to do, namely, predicting spending. A study conducted by the Society of Actuaries found that among 9 different prospective risk adjustment systems, none were able to predict more than 30% of the variation in spending across patients, and even when 6 of the risk adjustment systems were tested using concurrent information on patient conditions, the best system only predicted 55% of the variation and most predicted less than 40% of the variation in spending across patients.[122]

Ironically, the inability of risk adjustment systems to accurately predict spending is actually good news for efforts to reduce healthcare spending, because if the risk adjustment systems were 100% accurate in predicting spending based on patient health characteristics, it would mean that all differences in spending were directly related to differences in patient health conditions, which in turn would suggest that most services are appropriate and there is little opportunity to lower spending. Since many studies have shown that there is huge unjustified variation in the services different providers deliver and there is considerable overuse and underuse of many types of services, one should expect that services and spending would only be partially correlated with differences in patient needs. However, there is no way to know whether a risk score that is better at predicting variation in *total* spending across patients is doing a better job of predicting the *necessary* components or the *avoidable* components of the total variation. Consequently, it is essential

to do comparisons of spending for clinically similar groups of patients rather than relying on a single risk score.

### *Dealing with Outliers*

Examining patients in clinically similar categories can also reduce the need to make special provisions for "outlier" patients. Since no formula-based risk adjustment system can ever completely account for all differences among patients that could legitimately affect all of the services and the total spending they need, there will always be patients for whom actual spending is much higher than what a risk score would predict. The typical approach used is either to exclude these patients entirely from spending calculations or to arbitrarily truncate the expenditures on services they received.[123] However, this fails to distinguish whether the cause of the higher-than predicted spending for a patient is that the patient's needs are not adequately modeled by the risk adjustment formula or that the providers involved with the patient have used unnecessarily many or expensive services. By using clinical categories, the patient can be compared to similar patients using increasingly detailed subgroups to determine the extent to which the patient's characteristics are different from other patients and the extent to which the spending is different from patients with similar characteristics.[124]

*EXAMPLE: Assume that a group of patients with similar clinical characteristics typically need health care services costing $5,000 per year on average to treat their health conditions, but there is a one-in-one-thousand (.001) chance that a patient will have an unusual, additional condition that would increase the costs of treatment to $50,000. Assume further that the overall risk adjustment model does not include a measure of this unusual condition or that the model is not accurate in predicting the services needed to address the condition. If one provider treats 1,000 patients per year who have the same basic clinical characteristics, then that provider would expect to have one of the unusual patients every year; it would spend $50,000 for that patient and $5,000 for the remaining 999 patients, for an average of $5,045 per patient. However, if a second provider only has 100 patients per year with the same clinical characteristics, then over the course of 10 years, the provider would expect to have an unusual patient in one year but not in the other years. In the years in which the second provider has no unusual patients, its spending would be $5,000 per patient. However, in the year in which the second provider had the unusual patient, it would spend $50,000 for that patient and $5,000 for the remaining 99 patients, for an average of $5,450. This means that in the years in which the second provider had no unusual patients, the second provider would appear to be 0.9% less expensive than the first provider, but in the year in which it had the unusual patient, the second provider would appear to be 8% more expensive than the first provider. Over the course of ten years, both providers would have identical spending, but in individual years, they might appear very different due to random variation in patient characteristics. By identifying the patients with the unusual condition and examining their spending separately, it could be determined whether the provider was spending similar amounts (i.e., $50,000) on those patients as other providers, and whether the provider was spending similar amounts (i.e., $5,000) on the patients without that condition.*[125]

### *Creating Overall Provider Scores*

If it is desired, an overall "efficiency rating" of a provider can still be generated from the comparisons made within the disaggregated subcategories. The first step would be to divide the provider's spending for each clinically similar subgroup of patients by the average or median spending by other providers for patients who have the same characteristics. This series of ratios – some greater than 1.0 for patient subgroups where the provider's spending is higher than the average or median, and some less than 1.0 where spending is lower – could then be weighted based on either the relative number of patients in each subgroup for that provider, the relative amounts of average spending for each subgroup, or some combination of the two factors, and then the weighted ratios could be added to compute an overall score for that provider.[126] Subgroups where the provider did not have a sufficient number of patients to make meaningful comparisons could either be dropped or assumed to have a score of 1.0.[127] If the provider's overall score was greater than 1.0, it would mean that the provider had higher-than-average spending for the subsets of patients for which reliable comparisons can be made. Although this overall measure would be similar in appearance to what would be computed by dividing total spending by a single risk score, generating the measure based on category-by-category comparisons would be more statistically and clinically valid, and the measure could be directly broken down into measures for specific categories of similar patients to more clearly show individual providers where and how spending could be reduced for the types of patients and conditions they care for.

## Using Concurrent Risk Adjustment

Whether patients are disaggregated into categories using a categorical clinical model or individual risk models are developed, the patient categories or risk models should be based on complete information about patients' health problems that occurred during the time period in which spending is being measured, rather than only the kinds of historical information used in purely prospective risk adjustment systems. Not surprisingly, these "concurrent risk adjustment" approaches have been shown to be much more accurate in predicting expenditures than prospective systems.[128]

CMS and other payers have been concerned about problems that can arise in using concurrent risk adjustment as part of shared savings and similar payment systems that calculate how spending has changed over time as well as how it compares to other providers. Some form of risk adjustment is clearly needed in such payment models, because a provider's per patient spending can increase over time if the provider has new patients with more health problems or more severe problems as well as because the provider is delivering care less efficiently or effectively to existing patients. However, during the first year of these types of payment programs, the risk scores

for *all* patients will typically increase compared to the previous year, even if the patients have had no change in their health status, simply because the physicians in the new payment system now have an incentive to more completely and accurately record all of the patients' existing health problems.[129]

The fact that higher risk scores do not reflect a true increase in the patients' risk level, merely a change in the completeness and accuracy of the information about the patients, is not corrected by using a prospective risk measurement system based on out-of-date and incomplete information derived from the prior year. A better approach would be to not only determine the appropriate clinical categories for a patient based on all of the patient's health conditions in the *current* year, but to retroactively adjust the baseline risk scores/categories for the patients after the provider documents the existence of health conditions that existed prior to the current year. This would avoid crediting a provider with "savings" when spending increased simply because the risk score also increased as a result more complete documentation.

CMS is using what is essentially concurrent risk adjustment for some of the spending measures in its value-based purchasing programs for physicians and hospitals. The episode measures CMS has created are being individually risk adjusted based on the diagnosis information about the patient in the 90-180 days prior to the beginning of the episode,[130] and since most of the episodes are defined based on the patient's diagnoses during the episode, the result is that the spending is being adjusted by the most recent information about the patient's health problems.[131]

## Using Clinical Information from EHRs, HIEs, and Registries In Addition to Claims Data

As described in Section III, basing risk scores or patient categories solely on the information contained in claims data is problematic. Consequently, wherever possible, clinical data from electronic health records (EHRs), health information exchanges (HIEs), and clinical registries should be used in determining how a patient should be classified in addition to diagnosis codes recorded in claims data.[132]

No individual physician or provider's electronic health record and no specialty- or condition-specific registry will have complete information on all of a patient's health conditions or the procedures they have received. Even data available through a health information exchange will be limited to the providers who are participating in that exchange. There will be likely always be some information in claims that is not available from clinical data sources, and vice versa, so ideally the information on patient characteristics used for risk adjustment would come through both sources. Some provider groups and communities are developing the capability to merge information from claims and clinical data sources to create this more comprehensive source of information.

## Disaggregating by Non-Health Factors to Identify Impacts on Spending

As noted in Section III, there are important factors other than health conditions that can affect what services a patient needs. Instead of trying to incorporate these many different factors into an even more complex formula to produce a new version of a single risk-adjustment score, the disaggregation approach described earlier can be used to determine which of these factors may be influencing differences in spending in specific cases. For example, within any subgroup of patients defined by health conditions (e.g., "patients with multiple chronic conditions"), the patients could be further disaggregated by functional status, patient activation, health insurance status, socioeconomic status, etc. to see whether the level of services or avoidable complications differs significantly across those sub-categories. If they do, it would indicate that those additional factors had an important impact on spending after controlling for health conditions. Providers could then focus efforts on the subsets of patients where those factors were an issue.

For example, if high spending on emergency rooms for COPD patients primarily occurred among patients who were unable to drive or walk, the provider might arrange for home visits for those patients' care; if high spending primarily occurred in the subset of patients with low patient activation scores, the provider could focus education efforts on those patients. But if high spending on emergency room visits occurred for all categories of patients, the provider might need to find ways to improve access to the practice for all patients, such as more open appointments and better phone access.

*Disaggregating* spending into different categories of patients is also preferable to *adjusting* overall spending based on patient characteristics because it enables disparities between different groups to be measured and acted upon, rather than hidden inside a risk adjustment formula. For example, healthcare providers who care for large numbers of low-income patients express concern about having their spending and outcomes compared to providers who do not care for many low-income patients without adjusting for the difference in patient populations. However, patient advocates express concern about simply adjusting away the differences in spending and outcomes rather than trying to eliminate the differences. Disaggregating based on patient characteristics that should not, in an ideal world, affect spending or outcomes, and then comparing providers' performance on patients who are similar on these characteristics, allows the disparities in performance both within categories and between categories to be identified and acted upon.

Disaggregation has its limits, however, because the more characteristics that are used to define homogeneous patient groups, the fewer patients there will be in any group, and the less reliable comparisons among providers for patients in the groups will be. This makes it important to do analyses using multi-payer claims data wherever possible, so that the groups will be larger for analytic purposes than any individual payer could achieve solely with its own patient data. Analyses done using national databases, which would have even larger numbers of patients, could

help to identify which factors are the most important in identifying differences in patient needs, and then analyses done at the state or local level could focus disaggregated analyses just on those factors.

## D. Comparing Providers That Are Comparable

As discussed in detail in Section III-F, structural factors cause some providers to incur higher costs than others even if they are delivering similar services to similar groups of patients. Providers in rural areas, teaching hospitals, providers in communities with high costs of living, and providers with larger numbers of underinsured and uninsured patients may incur higher unit costs to deliver similar services to similar patients than other providers. If these differences are not addressed, they can distort comparisons between providers and between regions and unfairly penalize providers in regions where it is more expensive to deliver high-quality care.

Distinguishing differences in patient needs is an essential first step in addressing this issue, because there are typically also differences in the types of patients served and the types of services delivered by the providers in these different categories. For example, an academic medical center and its physicians will ordinarily treat more patients with unusual health problems than will a community hospital or community physicians, and these differences in patients will not be adequately addressed by typical risk adjustment systems for all of the reasons described in Section III. The clinical category approach described in Section IV-C will better address this by separating the services and spending delivered to patients with more common health problems from the services and spending delivered to the patients with rare conditions. (There may be too few of the patients with very rare conditions to allow statistically valid comparisons, but that is not solved by grouping them together with other patients.) This approach would help distinguish whether a hospital uses expensive services at a high rate because it has a large number of patients with conditions that require those services or simply because the hospital chooses to deliver those expensive services to patients whose needs could have been adequately addressed with less expensive services.

Once comparisons are being made based on comparable patients, two approaches can be used to address the differences in costs between different types of providers:

- **Adjusting spending for relative differences in costs.** To the extent that the impact of structural differences in costs can be quantified, a provider's spending could be adjusted upward or downward based on its relative costs before making comparisons with other providers. For example, if it were estimated that the teaching and research activities of a teaching hospital increase its costs by x% over an otherwise comparable non-teaching hospital, then spending at the teaching hospital could be reduced by x% when making comparisons to non-teaching hospitals.[133] This requires access to sufficient data to determine the extent to which cost differentials are caused by the structural factor being examined versus differences in the efficiency of the providers.
- **Comparing spending within peer groups.** An alternative approach is to compare providers only to similar providers, e.g., comparing teaching hospitals only to other teaching hospitals, comparing providers in rural areas only to other rural providers, etc. Similar to making comparisons for patients with similar needs, this has the advantage of making comparisons more straightforward and understandable than using an adjustment formula. However, if too many different factors are used to define a "peer" group, there may be too few providers in the peer group to allow statistically valid comparisons. Moreover, there would still need to be a way to determine whether the average difference in spending *between* peer groups was justified by structural differences outside of the providers' control.

Neither of these approaches will completely and accurately distinguish controllable vs. uncontrollable differences in costs, but they can improve the ability to identify providers who are performing significantly better or worse than others in comparisons of providers designed to identify differences due to the way care is delivered. However, even if the only reason a provider is more expensive is due to reasons unrelated to the efficiency or effectiveness of care (e.g., medical education costs or higher costs of charity care), the provider will still be more expensive from a patient's or payer's perspective if the structural differences in costs are financed through higher prices on individual services. To address this, communities need to develop ways of paying directly for the costs of medical education, charity care, etc. rather than forcing providers to finance them through higher prices on individual services.

## E. Addressing Differences in the Outcomes of Services and Spending

All of the methodologies and adjustments described previously in this report are focused narrowly on making fair and actionable comparisons of providers on spending. However, healthcare spending is obviously not an end in itself, nor should the goal be to drive spending as close to zero as possible. Healthcare spending is intended to achieve better health and a higher quality of life for patients, and higher spending on healthcare services should be welcome if it results in a sufficient improvement in patient outcomes.

Consequently, in addition to adjusting spending comparisons for differences in patient needs, it is also important to distinguish physicians, hospitals, and other providers that spend more and achieve better health outcomes from those providers that spend more but do not achieve better outcomes and from those providers that spend less but achieve significantly poorer outcomes.

Some people have suggested that a measure of "value" can be created simply by dividing a measure of quality by a measure of spending. Unfortunately, since quality and cost are measured on different scales, a higher ratio of quality to spending does not necessarily mean that a ser-

vice has higher value. For example, the ratio would increase if quality were reduced and spending were reduced by a larger amount, but that could reflect rationing of care rather than a true improvement in value. Similarly, the true value of a service could improve even when the ratio decreased if the smaller percentage increase in quality was viewed as "worth" the higher percentage increase in spending.

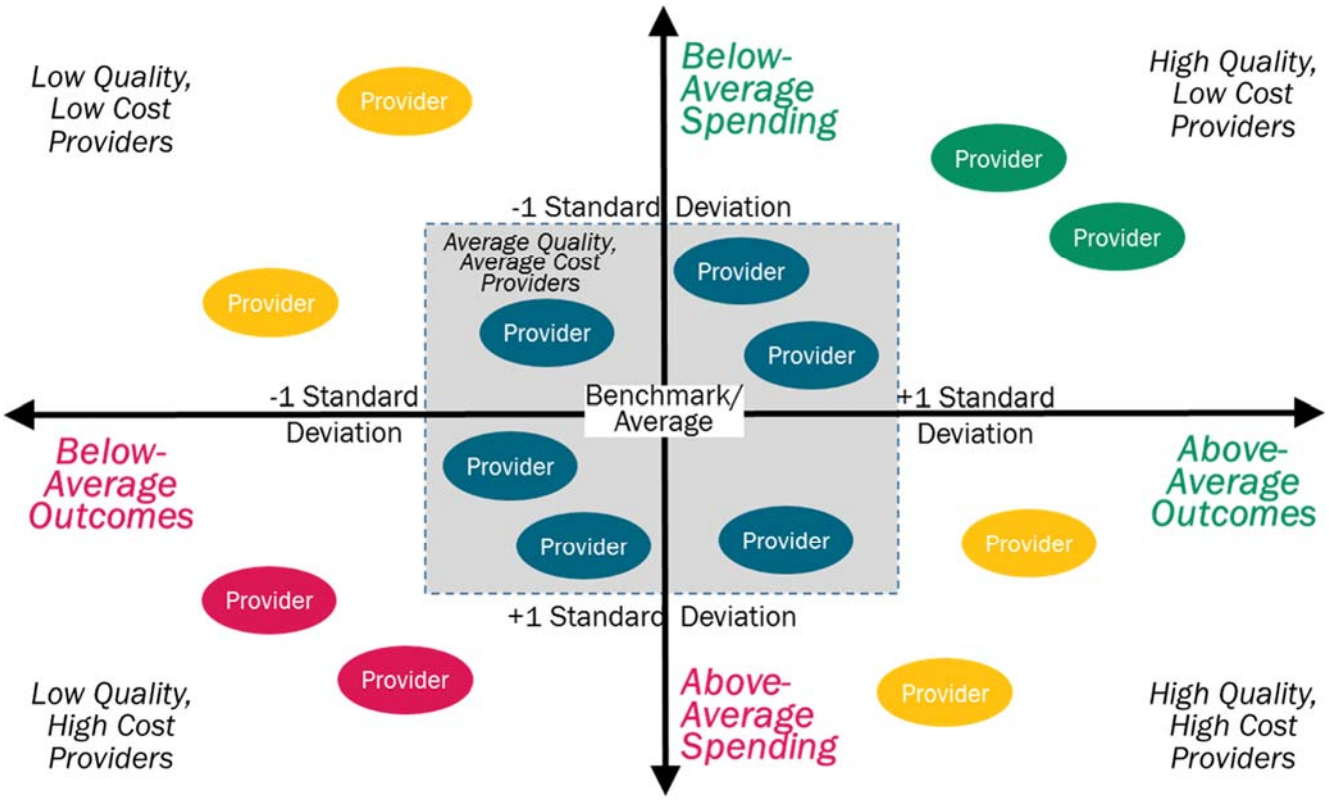
*EXAMPLE: Assume that Provider 1 delivers cancer treatment to a group of patients at a total cost of $25,000 per patient and Provider 2 delivers a different type of cancer treatment to patients with similar characteristics at a total cost of $50,000. The patients treated by Provider 1 live an average of 5 years, and patients treated by Provider 2 live an average of 8 years. Dividing the outcome by the cost shows that Provider 1 delivers 10.4 weeks of life per thousand dollars of treatment, while Provider 2 delivers only 8.3 weeks of life per thousand dollars of treatment. That would imply that Provider 1 is the higher-value provider, yet most people would likely say the opposite, since Provider 2 gives people an average of three extra years of life at an additional cost of only $25,000. If the two providers had different survival rates at the same treatment cost, or the same survival rates at different costs, the ratio wouldn't be needed to make the comparison, but when costs and outcomes both differ, the ratio is not very helpful in determining which provider is "better."*

Since health outcomes and spending are measured on different scales and since different people may place different dollar values on the same outcome, the best approach is to show how providers differ on both measures so that individual patients and payers can make their own judgments about which providers are "better."

As shown in Figure 20, a quality or outcome measure for providers would be compared to a benchmark (e.g., the average of all providers on that measure, or an absolute level of performance established in some way), and each provider's relative performance on the measure would be shown by their position along the horizontal axis. The providers' spending relative to a benchmark would be shown on the vertical axis; providers with lower spending ratios are shown above providers with higher spending ratios, so that "higher" on the axis means "better" (lower) spending. If a provider is in the upper right quadrant of the chart, it means they are better than benchmarks on both outcomes and cost; if they are in the lower left quadrant, it means they are worse on both measures.

Regardless of whether one is measuring quality or spending, there will be variations in spending from year to year and provider to provider based on variations in patient needs, differences in unit costs, etc. that cannot be captured adequately by any risk adjustment system or other means. A provider who has slightly better quality or spending scores than the benchmark this year may be slightly worse next year simply due to normal statistical

FIGURE 20
Comparing Providers on Both Spending and Outcomes



variation. Consequently, only providers whose performance is *significantly* better or worse compared to the benchmark (or compared to each other) should be flagged as higher or lower value. This is represented in Figure 20 by the gray box in the center – any provider whose spending or quality measure is in this box is not sufficiently different from the benchmark to declare them as being higher or lower value than others in the box.

With this kind of analysis, a provider who is (1) significantly better than another provider on one measure and (2) significantly better or not significantly different on the other measure can be viewed as delivering "higher value" care. So providers in the upper right corner are better than the providers in any of the other three quadrants, and providers in the upper left corner and lower right corner are better than those in the lower left quadrant. It is unclear, however, whether providers in the lower right corner and upper left corner are better or worse than each other, because each is better on one measure but worse on the other, and the tradeoff between cost and quality is a personal judgment. Different individuals might decide differently how to make that tradeoff; some might choose a provider with somewhat poorer outcomes in return for lower spending, while others might believe a provider's better outcomes justify the associated higher spending.

CMS is using this type of approach in its Value-Based Payment Modifier for physicians. Physicians are measured separately on quality and costs, and only physicians who are more than one standard deviation higher or lower than average are classified as having high or low spending or quality. This is far superior to "tournament" pay-for-performance systems that inappropriately rank-order providers on spending and quality measures and reward those providers that have even marginally higher performance than others, when the higher performance could simply be due to random variation or unmeasured differences in patient needs.

## F. Validity, Reliability, Comprehensiveness, and Actionability of Better Methods

Disaggregating spending into the categories and subcategories described above requires considerably more work than what is involved in current methods of reporting and assigning accountability for the total cost of care. Does it provide sufficiently better information to make it worth the effort?

### Validity

One of the most important criteria for evaluating a performance measure is its *validity*, i.e. whether it accurately captures the desired aspects of performance while excluding other, unrelated information. As the discussion in Section III makes clear, current approaches to assigning spending to providers do not produce valid measures of the decisions or actions by providers (since the measures include spending over which the providers have little or no influence) nor do they even produce valid measures of the total spending associated with the provider's own services, since many of the patients a provider sees will not be attributed to that provider.[134] They also do not produce valid measures of avoidable spending because of the weaknesses in risk adjustment and the lack of distinction between recommended, avoidable, and typical services.

The methodology defined in this section will produce more valid measures for accountability purposes by addressing all of these problems. The validity of *initial* reports produced using these methodology defined in this section may be lower than is desirable simply because of errors in how specific services are categorized resulting from incomplete or inaccurate coding of the claims data used to produce them. However, analogous problems exist with the data used to generate current spending measures. If the initial reports using the better-defined measures are produced to encourage and support action rather than used immediately for accountability purposes, these problems can be corrected so that future generations of the reports will be more accurate and more appropriate for use in accountability programs. More valid measures are more likely to be used by providers to support improvements in care, and when measures are used, errors are more likely to be found and corrected.[135]

### Reliability

As described in detail in Section III-G, measures of provider performance are commonly evaluated based on their *reliability*, i.e., whether the measure consistently distinguishes which of two providers is actually the better performer. True reliability is impossible to assess because there is no way to know for sure how much of the differences among providers are due to differences in the decisions that providers make for patients rather than differences in patient needs or random variation in services and costs. The cross-sectional reliability measures that are typically used can indicate the extent to which a spending measure is able to distinguish providers who had patients with high levels of spending in the past, but that does not mean the spending measure is a reliable way to tell patients which providers they should see in the future. Consequently, the significance of lower reliability will depend on whether the measure is being used to facilitate improvement or to reward or penalize performance.

There is no way to know how reliable the measures based on the categories defined in this section will be until they are actually generated and used. However, it seems likely that they would be at least as reliable as measures used today, if not more so, since they are designed to distinguish services that physicians control from services they do not control. If physicians deliver services in consistent ways to their own patients, but they deliver services differently from other physicians, then measures of spending that do a better job of distinguishing between what physicians do and order *themselves* from what *other* physicians do and order will inherently be more reliable. Moreover, if spending is compared on clinically similar groups of patients, rather than simply adjusted using a risk score, the comparisons will more reliably indicate how spending for particular types of patients would differ if they were receiving care from different physicians.

Making spending comparisons only for clinically similar groups of patients as recommended in Section IV-C may

reduce the statistical reliability of the measures because fewer patients will be compared in any subgroup than if the measures were simply adjusted based on risk scores and compared across all of the provider's patients. However, the statistically higher reliability of the aggregate measure would come at the expense of validity, since the single risk-adjustment score will not adequately address differences between patients with less common sets of health problems and will adjust spending for specialty providers based on patient characteristics that are not relevant to the care from that specialty. There is no value in having a system which reliably measures the wrong thing, so improving statistical reliability at the expense of validity accomplishes nothing.

## Comprehensiveness

Although the total spending for most patients will not be assigned to *any single* physician or provider under the methodology defined in this section, *all portions* of the spending on *every* patient will be assigned to *some* physician or other provider. As described in Section III, by attempting to assign all aspects of spending to a single physician or provider, the attribution methodologies currently being used by Medicare and other payers will leave *all* of the spending associated with *some* patients *unattributed* to *any* provider. Moreover, the true proportion of spending that is attributed using current methodologies is actually even lower than it appears because *some* of the spending for *many* patients is *inappropriately* attributed to physicians or other providers who had little or no influence over it. Although the current attribution methodologies could be changed to allocate a higher proportion of patients and spending to *some* provider by lowering the attribution threshold, the rate of *inappropriate* attribution would then increase as long as the methodology tries to assign a patient's total spending to a single physician or provider. The methodology described in this section is superior because it ensures that every aspect of spending is assigned to the provider best able to control it.

## Actionability

Finally, and perhaps most importantly, the methodology defined here is explicitly designed to facilitate action in reducing and controlling the growth in healthcare spending without harming patients. Rather than merely ranking providers on spending measures of questionable validity and reliability, providers would be given information on the services and aspects of spending they can control or influence, with a particular focus on types of services that are potentially avoidable or associated with health problems that could have been prevented, and that information would be presented in the context of relevant characteristics of their patients.

A more actionable report is more likely to result in reductions in healthcare spending in ways that do not harm patients. It is impossible to predict the actual return on investment because data in these formats are so rarely produced. However, since many experts believe that as much as 30-40% of healthcare spending is used for services that are unnecessary, inappropriate, or avoidable, if even a small reduction in that spending can be achieved through more actionable information, it could easily repay the investment needed to produce the more actionable information. Moreover, using the information to support the development and implementation of better payment systems, as described in Section VI, could result in an even greater return on investment.

# V. TAKING ACTION WITH MORE ACTIONABLE INFORMATION

The goal of better ways to analyze spending is not just to create fairer ways of holding healthcare providers accountable for what has happened in the past, but to give healthcare providers and all community stakeholders more effective assistance in reducing and controlling spending without harming patients in the future.

To see how the methodology and categories described in Section IV would do that, consider a hypothetical community that compiles and analyzes information on its healthcare spending using that approach. The first report it generates is at a community-wide level, an excerpt of which is shown in Figure 21.[136] The report shows that the community has 2% higher spending than other regions, and it reveals that two specific areas appear to be causing over 80% of the difference:

- For patients without chronic disease, there is much higher than average spending on potentially avoidable services, suggesting the possibility of overuse of these services in this community.
- For patients with multiple chronic diseases, there is much higher than average spending on potentially preventable conditions.

## FIGURE 21
### Differences in Spending Per Patient in a Hypothetical Community Compared to Other Communities

| Patient Categories | % of Total Patients | Spending Categories: Recommended Services | Potentially Avoidable Services | Potentially Preventable Conditions | Other Services | Total Spending |
|---|---|---|---|---|---|---|
| | | *Percentage Difference in Spending From Average in Other Regions* | | | | |
| Patients with 0 chronic diseases | 60% | -3% | +20% | +2% | +1% | +5% |
| Patients with 1 chronic disease | 20% | +2% | +1% | +3% | +1% | +4% |
| Patients with 2 chronic diseases | 15% | -1% | -1% | +10% | -1% | -1% |
| Patients with 3+ chronic diseases | 5% | -4% | 0% | +15% | +1% | +3% |
| All Patients | 100% | -4% | +3% | +7% | +1% | +2% |
| | | *Difference in Spending Per Patient Per Month From Average in Other Regions* | | | | |
| Patients with 0 chronic diseases | 60% | -$0.17 | $2.71 | $0.07 | $0.33 | $2.94 |
| Patients with 1 chronic disease | 20% | $0.62 | $1.04 | $0.48 | $1.58 | $3.71 |
| Patients with 2 chronic diseases | 15% | -$0.81 | -$2.24 | $5.37 | -$4.55 | -$2.24 |
| Patients with 3+ chronic diseases | 5% | -$10.22 | $0.00 | $42.74 | $15.76 | $48.28 |
| All Patients | 100% | -$0.61 | $1.50 | $3.08 | $0.62 | $4.58 |
| | | *Difference in Total Annual Spending for 10,000 Patients* | | | | |
| Patients with 0 chronic diseases | 60% | -$12,031 | $195,120 | $4,680 | $23,998 | $211,766 |
| Patients with 1 chronic disease | 20% | $14,784 | $24,852 | $11,599 | $37,810 | $89,045 |
| Patients with 2 chronic diseases | 15% | -$14,669 | -$40,408 | $96,624 | -$81,948 | -$40,401 |
| Patients with 3+ chronic diseases | 5% | -$61,319 | $0 | $256,437 | $94,583 | $289,702 |
| All Patients | 100% | -$73,235 | $179,564 | $369,340 | $74,442 | $550,112 |

Provider-specific reports in these two categories, similar to what are shown in Figure 22, indicate that the potentially avoidable services for non-chronic disease patients seem to be concentrated in Providers 1 and 2, but that spending on potentially preventable conditions for the chronic disease patients is common to all provider organizations.

A further breakdown of the spending by Providers 1 and 2 on the patients without chronic disease, shown in Figure 23, indicates that one of the biggest areas where overuse may be occurring is for treatment of patients with low back pain.

Not surprisingly, the high spending in the potentially avoidable services category for back pain is primarily due to high levels of spending on MRIs ordered for patients. However, further analysis shows that the high spending is not solely due to frequent use of MRIs; Figure 24, which disaggregates differences in spending into differences in utilization versus differences in price, shows that physicians associated with Provider 1 are ordering imaging at higher rates, but the higher spending for Provider 2 is caused primarily by patients who have imaging delivered at high-priced imaging facilities.

Turning attention to the patients who have chronic disease, Figure 25 shows that spending on potentially preventable conditions for patients with multiple chronic diseases stems from high rates of visits to emergency rooms for exacerbations of chronic disease and the hospitalizations which follow those visits. Discussions with all of the providers who are caring for these patients indicate that because they can only be paid for office visits, the physicians do not have the resources to provide better educa-

## FIGURE 22
### Differences in Potentially Avoidable Spending for Different Providers

| Patient Categories | Provider Organization | Potentially Avoidable Services | Potentially Preventable Conditions |
|---|---|---|---|
| Patients with 0 chronic diseases | Patients of Provider #1 | +20% | +3% |
| Patients with 0 chronic diseases | Patients of Provider #2 | +35% | -1% |
| Patients with 0 chronic diseases | Patients of Provider #3 | +9% | +5% |
| Patients with 0 chronic diseases | Patients of Provider #4 | -1% | -2% |
| Patients with 0 chronic diseases | All Patients in Community | +20% | +2% |
| Patients with 1 chronic disease | Patients of Provider #1 | +4% | +1% |
| Patients with 1 chronic disease | Patients of Provider #2 | -3% | +5% |
| Patients with 1 chronic disease | Patients of Provider #3 | +2% | +4% |
| Patients with 1 chronic disease | Patients of Provider #4 | -1% | +6% |
| Patients with 1 chronic disease | All Patients in Community | +1% | +3% |
| Patients with 2 chronic diseases | Patients of Provider #1 | 0% | +10% |
| Patients with 2 chronic diseases | Patients of Provider #2 | -2% | +8% |
| Patients with 2 chronic diseases | Patients of Provider #3 | -2% | +11% |
| Patients with 2 chronic diseases | Patients of Provider #4 | +3% | +13% |
| Patients with 2 chronic diseases | All Patients in Community | -1% | +10% |
| Patients with 3 chronic diseases | Patients of Provider #1 | +4% | +11% |
| Patients with 3 chronic diseases | Patients of Provider #2 | +3% | +13% |
| Patients with 3 chronic diseases | Patients of Provider #3 | -4% | +17% |
| Patients with 3 chronic diseases | Patients of Provider #4 | -1% | +19% |
| Patients with 3 chronic diseases | All Patients in Community | 0% | +15% |

## FIGURE 23
### Differences in Spending From Other Communities on Potentially Avoidable Services for Patients With No Chronic Diseases

| Provider | Diagnoses Reported for Services | Services Performed by Provider | Services Ordered by Provider | Services From Referrals by Provider | Related Services | Unrelated Services | Total |
|---|---|---|---|---|---|---|---|
| Provider #1 | Dizziness | +1% | +5% | +10% | +7% | 0% | +9% |
| Provider #1 | Lower Back Pain | -1% | +60% | +10% | +5% | +3% | +35% |
| Provider #1 | Upper Respiratory Infection | +1% | +7% | 0% | 0% | 0% | +5% |
| Provider #1 | Other | -2% | +3% | +5% | +4% | -1% | +4% |
| Provider #1 | Total Patients | +0% | +40% | +9% | +5% | +2% | +20% |
| Provider #2 | Dizziness | +2% | +10% | +15% | +12% | +3% | +11% |
| Provider #2 | Lower Back Pain | +5% | +80% | +15% | +20% | +4% | +60% |
| Provider #2 | Upper Respiratory Infection | +3% | +15% | +5% | +6% | -1% | +10% |
| Provider #2 | Other | +1% | +5% | +12% | +10% | +2% | +9% |
| Provider #2 | Total Patients | +3% | +40% | +16% | +13% | +2% | +35% |

## FIGURE 24
### Differences From Other Communities in Utilization and Unit Cost of Imaging Ordered for Patients With Lower Back Pain

| Provider | Diagnosis Reported for Services | Service Ordered | Services Ordered Per 100 Patients with Diagnosis | Actual Cost Relative to Typical Cost | Total Spending on Service |
|---|---|---|---|---|---|
| Provider #1 | Lower Back Pain | MRI | +70% | -6% | +60% |
| Provider #2 | Lower Back Pain | MRI | +5% | +71% | +80% |

tion and self-management support for the patients. The providers and payers in the community work together to develop a plan for paying for additional patient support services, and the information in the spending report is used to develop a business case[137] showing how total spending can be reduced even with the additional investment in services through the kind of reduction in the frequency of emergency room visits that has been achieved in demonstration projects around the country. Because all physician practices need to provide additional patient support services, but most of the physician practices are small and each has relatively few patients who need the services, the physician practices develop a community-wide mechanism for sharing patient support services, so that the overall cost of delivering those services would be lower; the payments from the payers could then be used to pay each practice's share of the total cost of the new services.

In the months ahead, the community can produce these same detailed breakdowns on a regular basis in order to monitor progress in reducing the identified areas of overspending. The overall reports on spending should also continue to be produced in order to identify additional or new areas of opportunity for reducing spending.

## FIGURE 25

Causes of High Levels of Potentially Preventable Spending for Patients with Chronic Disease

| *Patients with 3+ Chronic Diseases* | Difference in Per Patient Spending From Other Communities | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | ER Visits for Potentially Preventable Conditions | | Hospital Admissions for Potentially Preventable Conditions | | Hospital Readmissions for Potentially Preventable Conditions | | Other Services Associated with Potentially Preventable Conditions | Total Spending on Potentially Preventable Conditions |
| Provider Organization | Due to Exacer-bations of Chronic Disease | Due to Other Issues | Due to Exacer-bations of Chronic Disease | Due to Other Issues | Due to Exacer-bations of Chronic Disease | Due to Other Issues | | |
| Provider #1 | +42% | +4% | +37% | +7% | +21% | -3% | +9% | +11% |
| Provider #2 | +40% | -3% | +30% | -1% | +17% | +4% | +6% | +13% |
| Provider #3 | +25% | +6% | +27% | -2% | +12% | -3% | +3% | +17% |
| Provider #4 | +30% | +2% | +20% | +10% | +15% | +1% | -2% | +19% |
| All Patients in Community | +34% | +2% | +28% | 4% | +16% | 0% | +4% | +15% |

# VI. MOVING FROM MEASUREMENT TO ACCOUNTABLE PAYMENT

Developing *actionable information* on healthcare spending is a critical first step in trying to reduce or control healthcare spending. As demonstrated in the example in Section V, the methodology described in Section IV would better enable the identification of which patients, which providers, and which services may be opportunities for reducing spending. However, lack of actionable information is not the only barrier to controlling healthcare spending; the payment system is also a major barrier.

## A. Better Ways to Pay for Healthcare

Current fee-for-service payment systems create several significant barriers to implementing the changes in care delivery that would reduce spending without harming patients:

- **Some services that could lower overall spending aren't paid for adequately or at all.** For example, Medicare and most health plans don't pay physicians to respond to a patient phone call about a symptom or problem, even though those phone calls can avoid far more expensive visits to the emergency room. Medicare and most health plans won't pay primary care physicians and specialists to coordinate care by telephone or email, yet they will pay for duplicate tests and the problems caused by conflicting medications. A physician practice that does outreach to high-risk patients or hires staff to provide patient education and self-management support typically can't be reimbursed for those costs, even if the services help avoid expensive hospitalizations or allow diseases to be identified and treated at earlier stages.
- **Physicians, hospitals, and other healthcare providers are often financially penalized for reducing unnecessary services and improving quality.** Under the fee for service system, providers lose revenue if they perform fewer procedures or lower-cost procedures, but their costs of delivering the remaining services generally do not decrease proportionately,[138] which can cause operating losses for the providers. Most fundamentally, under the fee for service system, providers don't get paid at all when their patients stay well.

These barriers cannot be solved by merely adding bonuses or penalties based on healthcare spending measures on top of the *current* payment system. A small pay-for-performance bonus may not generate enough revenue to pay for services that are not paid for adequately in the current fee-for-service system or to offset the financial penalties providers currently face in reducing unnecessary services. Moreover, if the measures used for the bonuses and penalties have the flaws described in Section III, they can create perverse incentives for providers to avoid caring for patients who could benefit the most from improved care.

Instead, *different* payment systems are needed to truly overcome the barriers. Five types of payment reforms can be used to overcome these barriers[139]:

- **Bundled payment**, i.e., a single payment for all components of a service delivered by all providers, regardless of the setting. A bundled payment gives providers the flexibility to deliver service components that are not reimbursed now and to redesign the overall service without worrying about exactly what service components will be reimbursed.
- **Warrantied payment**, i.e., a higher payment for high-quality delivery of a procedure or service, with no payment for any additional services needed to correct errors and avoidable complications resulting from the original procedure. A warrantied payment gives a provider the upfront resources needed to redesign care and enables the provider to generate higher margins by delivering higher-quality care.
- **Episode payment for a procedure**, i.e., a single payment for all services associated with a procedure during a specific period of time after the procedure occurs, including services to correct preventable complications. An episode payment for a procedure gives providers additional flexibility to redesign care and reduce complications, not just with respect to the procedure itself, but also with respect to follow-on services.
- **Condition-based payment**, i.e., a single amount for all services and procedures needed to treat a particular patient condition or combination of conditions. A condition-based payment gives providers the flexibility to use different procedures or services to treat a patient's condition without worrying about incurring losses if fewer services or procedures are performed or if procedures are performed in lower-cost settings.
- **Global payment**, i.e., a risk-adjusted payment for all services and procedures needed to treat all of the health problems for a group of patients. Global payment is the most *flexible* payment of all, enabling providers to target additional resources on conditions where spending could be reduced and to coordinate care among the multiple providers dealing with patients with multiple health conditions. Global payment is also the most *accountable* form of payment, since it requires providers to manage spending on *all* of the services the patients need.

## B. How Actionable Spending Information Can Support Payment Reform

One of the biggest barriers to implementing these types of improved payment systems is the difficulty providers face in getting the kind of information that is needed to develop the specifics of the payment systems and to set appropriate prices. The structure for aggregating and disaggregating data that is described in Section IV provides data in exactly the format that is needed for defining and pricing better payment systems.

- The distinctions made in Section IV-A between services delivered or ordered by the provider versus those delivered or ordered by other providers, and the distinctions that are also made there between the number of services, the expected costs of those services, and the actual prices of those services, are essential for developing successful payment systems. If a physician, hospital, or other provider cannot control or reasonably expect to influence the utilization or price of services, then they will not be willing to accept a payment system that holds them accountable for spending on those services. Conversely, if a payment system is designed to give a provider accountability for aspects of spending they *can* control or influence as well as the flexibility to *change* the way services are delivered in order to impact that spending, then providers, purchasers, and patients all have the potential to benefit. Spending Categories 1, 2, 3, and 4 help define which services should be included in bundles, warranties, episodes, and condition-based payments, and which should be excluded.
- Another concern that healthcare providers have about participating in new payment systems is that they will receive the same amount of payment regardless of how severe or complex their patients' needs are. The methods of categorizing patients described in Section IV-C can mitigate or eliminate these concerns, since they allow higher payment amounts to be defined for patients with greater needs.
- In order to set an appropriate price for an accountable payment, a provider needs to know where services and spending can be reduced and by how much. The actionable information defined in Section IV-B enables a provider to identify potentially avoidable spending as well as areas of underuse where spending may need to increase.

### Defining and Implementing Bundled Payments

The simplest payment reform is a "bundled" payment, i.e., making a *single* payment for multiple services that are currently paid for *separately*. The most logical services for a provider to include in a bundled payment would be those identified in Spending Categories 1, 2, and 3.

- **Bundling Services in Spending Category 1.** If analyses of the data in Spending Category 1 show that providers use different combinations of fee-based services in addressing similar patient needs, then rather than paying separately for those services, a single bundled payment or "case rate" could be defined to cover all of them. The bundled payment would give the provider greater flexibility to change the number of services based on the patient's needs while providing a more predictable payment for both the provider and the payer. For example, since 1983, hospitals have been paid by Medicare on a bundled payment basis for the services they deliver during inpatient care, and they have received bundled payments for outpatient services since 2000. Some physicians are paid on a case rate basis, such as the "global fee" paid to obstetricians for care of pregnant women that bundles prenatal care services and the delivery of the baby into a single payment.

  Comparisons among providers of the total spending on services in Category 1 for patients with similar clinical characteristics would help determine the appropriate size of the bundled payment. Those comparisons would also help providers with higher per-patient spending on the bundled services determine how patient care could be redesigned to use fewer services than are being delivered today. Comparisons across different clinical categories could be used to determine if different numbers of services are needed for different types of patients, and if so, how the bundled payment amounts should differ for patients with different characteristics.
- **Bundling Services in Spending Categories 1 and 2.** If analyses of the data in Spending Category 2 show that providers routinely deliver a service in conjunction with services from other providers in addressing a particular set of patient needs, or if some providers use services from other providers to do so and others do not, a single bundled payment could be defined that would include all of these services and providers. This would give the providers the flexibility to redesign the care in order to use fewer services or lower-cost services while providing the payer with a more predictable payment. Medicare has done this in its Acute Care Episode Demonstration program, combining the hospital and physician payments for a number of cardiovascular and orthopedic procedures. The total payment by Medicare is lower than the previously separate payments, and the hospital and physician can then work together to identify ways to reduce the costs of delivering the care.[140]
- **Bundling Services in Spending Categories 1, 2, and 3.** If analyses of the data in Spending Category 3 show that providers order different numbers and types of services from other providers in addressing similar patient needs, a single bundled payment could be defined that includes both the services delivered directly by the provider and these ordered services. The provider would then have the flexibility to order different sets of services based on patient needs with the accountability for ensuring that, on average, the costs of the services delivered and ordered stayed within the bundled payment amount. In return, the payer would have a more predictable payment for all of the services that is lower on average than its current spending. The amount of spending in Subcategories 1(b), 2(b), and 3(b) – the potentially avoidable services – would help in making estimates of how much lower the bundled payment could be than current

spending and would help providers with high spending identify ways to reduce unnecessary services. Conversely, if spending in Subcategories 1(a), 2(a), and 3 (a) are low, there may be need to be increases in spending in those areas to correct underuse of recommended services.

## Defining and Implementing Warrantied Payments

If a significant amount of spending for particular types of providers is associated with services for potentially preventable conditions (this spending would be included in subcategories 1(c) and 4(c) for those providers), then a warrantied payment could be defined. The warrantied payment would pay more for recommended and typical services (i.e., the spending included in subcategories 1(a) and 1(d)), while eliminating payment for some or all of the services associated with the potentially preventable conditions, such that the total payments to providers with *low* rates of potentially preventable conditions would be higher than today, but *total* spending by payers would also be lower than today.

*EXAMPLE: A hospital and a physician want to offer a procedure with a warranty for readmissions, i.e., the hospital and physician would be paid more for the procedure than they are today, but they would not be paid at all if a readmission occurred that was related to the procedure. The hospital and physician determine how much is currently being spent on readmissions for patients who receive the procedure, they estimate how much they could reduce the readmission rate, and they project how much more they would need to spend, if anything, to achieve that reduction. Comparisons to other providers on spending would help in making these estimates and projections. The hospital and physician propose a price for the warrantied procedure that is sufficient to cover the cost of performing the procedure, the cost of any additional services designed to reduce readmissions, and the cost of treating the (smaller number of) patients who would be readmitted. If the payer determines that the proposed price is lower than what it is currently spending for patients who receive the procedure (including the average spending on patients who are readmitted), then the payer would save money by contracting with the hospital and physician on the proposed basis. If the current spending amounts were higher for patients with certain comorbidities and other factors that increase the probability of readmissions, then a higher payment could be established for those patients, but still with the warranty for readmissions. The specifics of the warranty (e.g., which specific types of readmissions would be covered by the warranty) would be defined in the contract with the payer.*

## Defining and Implementing Episode Based Payments for Procedures

An episode payment for a procedure could be defined using a combination of bundles and warranties.[141] Information needed to determine the amount of the payment could be derived from the spending tabulations for the physicians or other key providers who deliver the procedure. For the providers who deliver the procedure, their services associated with the procedure and with any preventable complications would be included in Spending Category 1. Any integrally related services delivered by other providers would be included in Spending Category 2. Services ordered by the provider who delivered the procedure would be included in Spending Category 3, and any related services, including preventable complications, would be included in Spending Category 4.

For example, for a surgery performed at a hospital, a single episode payment could be defined to cover the costs of the surgeon's services, the anesthesiologist's services, the hospital's services, the post-acute care services, and any related readmissions. The Geisinger Health System has done this through its ProvenCare program,[142] and CMS is implementing this through its Bundled Payments for Care Improvement initiative.[143]

## Defining and Implementing Condition-Based Payments

If there are different procedures that can be used to treat a particular patient condition, then a condition-based payment could be appropriate instead of or in addition to separate procedure-based episode payments. The grouping of patients into clinically similar categories as defined in Section IV-C enables the identification of opportunities for condition-based payment. If the types and mix of procedures for patients with a particular condition or combination of conditions varies significantly across providers (including different types of providers managing the same type of condition), then rather than paying separately for different procedures, a single condition-based payment would give a provider or group of providers caring for the patient greater flexibility to use the best procedure based on the patient's needs while providing a more predictable payment for both the provider and the payer. By examining services, spending, and outcomes for different subgroups of patients, it could be determined whether different payment amounts would be appropriate for different subgroups of patients who have the same condition but who also have other characteristics that make them more likely to require more services or more expensive services.

*EXAMPLE: Many pregnant women deliver their babies through Cesarean sections when vaginal deliveries would be better for the mother and baby as well as costing less. A condition-based payment could be defined to pay for pregnancy (rather than paying for a specific type of delivery), with the payment amount calculated based on the appropriate mix of vaginal deliveries and C-sections and their associated costs. The condition-based payment might only be applied to mothers without other health problems, or it could be risk-stratified to pay a higher amount for care of mothers with characteristics that would make them more likely to need the more expensive C-Sections or other services.*

For patients with chronic diseases, a condition-based payment could be designed to cover the full range of services the patients need from all providers in order to successfully manage their chronic conditions over the course of a

year. Moreover, rather than only making such payments for patients with single chronic diseases, condition-based payment amounts could also be defined for patients with multiple conditions. The analysis to support this would be based on the services currently being provided to patients with different combinations of chronic disease as determined from the tabulations of spending for patients in the clinically-defined categories defined in Section IV-C.

### Defining and Implementing Global Payments

The same approach described for condition-based payments can be used to define a global payment for a population of patients. Whereas a condition-based payment would only be designed to cover the services relevant to that condition (or combination of conditions) for patients who have the condition(s), a global payment would be designed to cover *all* of the services needed for *all* of the conditions the patients have. The global payment would need to be risk-adjusted, i.e., the amount of the payment would need to be higher for a group of patients who have more health conditions or more serious health conditions, since they would need more services. However, rather than simply calculating risk scores for the patients using a formula and then adjusting the overall payment based on the average score, the appropriate payment amounts could be defined for groups of patients in clinically defined categories (as would be done in defining condition-based payments), and then the overall global payment would be determined by multiplying the number of patients in each category by the appropriate payment amount for that category and summing those products to determine an appropriate overall global payment amount.

## C. Transitioning to Better Payment and Compensation Systems

The most appropriate type of payment reform depends on what the data in the spending analyses show are the opportunities for reducing spending. For example, if there is significant variation in the number of different services that providers use to deliver the same procedure, then a bundled payment for that procedure should be considered. If there is no such variation, then there may be little benefit to the effort needed to create a bundled payment. If different types of procedures are being used for similar patients, then a condition-based payment could be appropriate. If there is only one approach to treating a condition but a large number of potentially preventable complications are occurring with that treatment, then an episode-based payment with a warranty for that procedure might be more appropriate.

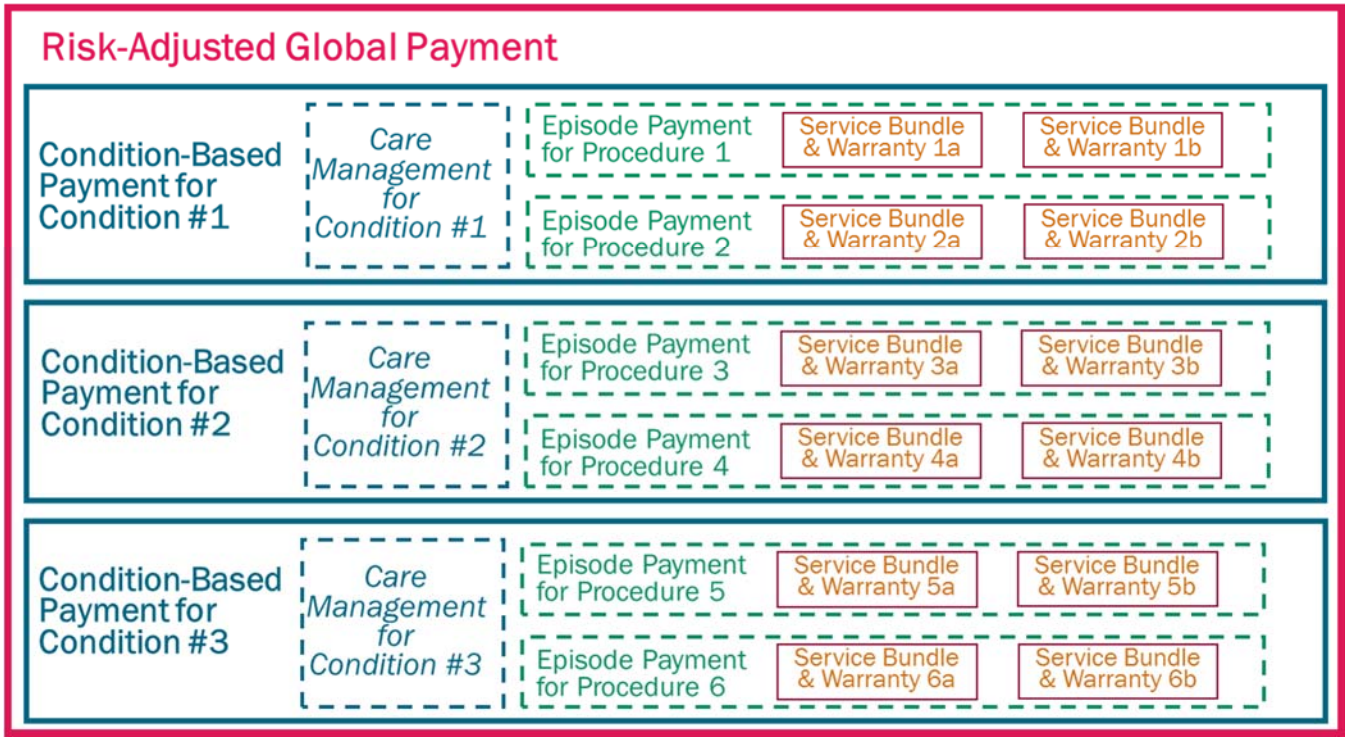
The payment models are not mutually exclusive, and indeed, they may all be appropriate for providers trying to implement global payments. If a multi-specialty physician group, health system, or accountable care organization received a risk-adjusted global payment for a population of patients, it could then use that to pay individual specialists using condition-based payments for managing specific types of health conditions those patients experience. When a patient with a particular condition needed a procedure, an episode-based payment could be paid to the providers performing that procedure from the overall condition-based payment. The episode payment, in turn, could be used to make bundled payments to subgroups of providers who deliver specific services within the overall episode, and each of those payments could include a warranty focused on the specific types of preventable conditions that each provider could prevent. These subsidiary payments could either be paid on a contractual basis to independent providers or used as the basis for an internal compensation structure for employed physicians and other providers. As shown in Figure 26, the five different payment models could be "nested," similar to a Russian Matryoshka doll.

Conversely, physicians, hospitals, and other providers could transition to global payments by starting with bundled payments and warranties within individual specialties, then combine the bundles and warranties into episode payments for procedures, then join forces with other providers who deliver different procedures to accept a condition-based payment for the health condition with which they are all involved, and then ultimately work with all of the providers who manage different types of conditions in order to accept an overall global payment for a population of patients.

These transitions and internal payment/compensation structures can work most effectively if all of the payments are defined in a mutually consistent way, using information on services and spending analyzed in the fashion defined in Section IV.

FIGURE 26

Using Bundled Payments, Warranties, and Episode Payments to Support Condition-Based Payments and Global Payments



# VII. OBTAINING THE DATA NEEDED FOR IMPLEMENTATION

As shown in Sections V and VI, data that are effectively organized to produce actionable information can both identify opportunities to reduce spending and support the development of payment reforms that enable providers to capitalize on those opportunities. Consequently, it is imperative to make this kind of information more widely available. Several things are needed to rapidly and efficiently develop the kinds of actionable information defined in this report so that this information can be used to redesign both care delivery and payment systems.

## Access to all-payer claims data on the services patients are receiving and the amounts being spent on those services

If they wished, individual payers – Medicare, commercial health plans, and state Medicaid agencies – could each use their own claims data to generate reports structured in ways similar to what is described in this report. However, if every physician and hospital receives information separately from each payer, the reliability of the information will inherently be lower because the number of patients in any category will be smaller than if all payers pooled their data. Moreover, if each payer develops its own reports, it is highly likely that the reports will be generated using different definitions of the categories, different risk adjustment systems, and different comparison groups. Consequently, if each payer generates separate and different reports, it will be more difficult for healthcare providers to review and act on this information.

Fortunately, there are now a growing number of states and communities where a state agency or a Regional Health Improvement Collaborative (RHIC) organization receives claims data from all or most of the payers in the community and uses it to provide information on the quality and cost of care to providers, payers and purchasers, and patients.[144] A provision in the federal Affordable Care Act has enabled a number of these organizations to become "Qualified Entities" and receive Medicare claims data to complement claims data from commercial health insurance plans and Medicaid programs.[145] These agencies and RHICs can generate more robust analyses of spending by combining data from all payers in the community, while also generating payer-specific analyses in a common format for all payers. Moreover, by working together across regions to use standard definitions, risk adjustment systems, etc., RHICs can also develop analyses that can support national policy decisions about payment systems.

Unfortunately, many of the communities with multi-payer claims databases only have access to information on the type and number of services provided, not on the *amount of money paid* for those services.[146] Since a number of research studies have shown that there is significant variation in the amounts that are paid to different providers for ostensibly the same services, both within geographic regions and across regions, it is impossible to produce accurate and actionable analyses of what is driving *spending* without access to information on payment amounts.

In addition, it is essential that organizations with access to multi-payer data have the flexibility to creatively *analyze* the data, not just to produce standardized measures. Data use agreements that are narrowly defined to only allow production of specific quality and spending measures provide no ability to produce the types of actionable analysis that this report has shown are essential for successful efforts to reduce spending and to move to payment systems that support better care.

## Access to clinical data

Although the analyses and comparisons described in previous sections can be generated using just the information on diagnoses, services, and payment amounts that are available in claims data files, these analyses will be far more accurate if they can draw on more complete clinical information about patients and the services they receive, such as the data available in electronic health records (EHRs) and patient registries. Some payers collect a limited number of elements of clinical data from providers, but others do not. A growing number of communities and medical societies have created patient registries containing clinical data that could be used to improve analyses of spending.

## Collection of data on patient outcomes

As noted in Section IV-E, the goal of healthcare reform should be to control or reduce spending while maintaining or improving the quality of care for patients. This can only be done by measuring the quality of care and the benefits to patients as well as the amount spent to deliver those results. There is growing recognition that traditional process measures of quality are insufficient and potentially counterproductive because they are not always closely correlated with outcomes and they may inhibit new innovations and patient-centered care. In most cases, measures of the true outcomes of care, e.g., patients' quality of life and productivity, are not even captured in clinical data records, much less in claims data. The only way to obtain this information is directly from patients. Patients are now being routinely surveyed about their experience of care from individual providers, and patients should also be surveyed to collect information about the outcomes they experience.

## Resources to generate better analyses of data and ensure they are used

Merely *obtaining* claims, clinical, and outcomes data is insufficient; the data must be combined and analyzed using the methodologies described in this report in order to support action. Although individual providers and payers might prefer to do all analyses on their own, it would be more economical if the analyses were done at the community level; moreover, the data would be more likely to lead to appropriate payment reforms if both providers and payers could trust and use the same set of information.

Multi-stakeholder Regional Health Improvement Collaboratives are an ideal mechanism for providing objective data and analysis that all stakeholders – purchasers, providers, and patients – can trust and use to help forge feasible strategies for controlling healthcare spending while maintaining and improving quality. However, adequate resources will be needed to support the work involved in assembling a full range of data, producing high-quality analyses, and helping stakeholders use the data to support action.

All stakeholders – the federal government, state government, health plans, employers, hospitals, physician groups, and consumers – will need to contribute funding both to ensure that sufficient resources are available in every community and also to ensure that all analyses can be produced in a neutral and transparent way. Since all stakeholders will benefit if healthcare spending can be successfully controlled in appropriate ways, this is one of the best investments they can make.

# VIII. CONCLUSION

High and growing healthcare spending is one of the most serious problems facing the United States and many other countries. Changing the way care is delivered in order to reduce spending without harming patients must be a high priority for everyone, both inside and outside of the health care industry.

Success will depend on active engagement and strong leadership from physicians, hospitals, and other providers. They collectively deliver the care to patients that currently costs too much and achieves too little in terms of quality and outcomes, and only they have the ability to change care in ways that can reduce spending without harming patients. However, their success in identifying opportunities to improve care will depend on their ability to obtain more actionable analyses of services and spending than they can routinely receive today. Moreover, their success in implementing the necessary changes in care will require payment reforms that remove the significant barriers to higher quality, lower cost care that exist in most current payment systems.

Considerable investments of time and money are being made to measure spending, attribute it to providers, and adjust providers' payments in an effort to encourage providers to reduce healthcare spending. If measurement, attribution, and accountability systems are not designed properly, they will not only fail to provide the actionable information providers need, they can discourage providers from making feasible changes by demanding they control services and spending that are beyond their range of influence. Pay for performance and shared savings programs based on spending measures not only fail to resolve the barriers created by fee-for-service payment, they can further discourage action by penalizing providers based on flawed systems of assigning accountability.[147]

Fortunately, as detailed in this report, there are better ways to analyze spending that can help physicians, hospitals, and other providers identify opportunities to achieve better outcomes at lower costs. There are also better ways to pay providers that will enable them to redesign care to implement those opportunities in ways that are financially feasible for them. Although significant investments of time and money will be needed to create better analyses of spending and to design and implement better payment systems, some of the resources needed for this can be shifted from existing, less effective measurement and payment systems, and the rest can be more than repaid from the savings in healthcare spending achieved through more effective measurement and payment systems.

The most successful way to implement these changes is likely to be through state and regional multi-stakeholder approaches, rather than a one-size-fits-all national approach. Many states and regions already have laid the foundation for this through initiatives to assemble data, create analytic capability, pursue innovative payment reforms, and facilitate collaboration among payers, providers, and other stakeholders. These communities can lead the way for the rest of the country if they receive the necessary support to do so.

# ENDNOTES

1. In this report, the term "provider" will refer to any individual or organization that delivers health care services to patients. This can include a physician, a nurse practitioner, a physician assistant, a physician practice, a hospital, a home health agency, and any number of other types of organizations. Because of the central role that physicians play in ordering and delivering services, most of the examples in the report will be focused on physicians, but most of the recommendations regarding analyses of spending are applicable where the provider is an entity other than a physician or physician practice, such as a hospital, as well as to a physician. In addition, the use of the term "provider" in this report can mean either an individual physician, multiple physicians, or physicians and other providers working together in a group practice, an Independent Practice Association, a Physician-Hospital Organization, a health system, an Accountable Care Organization, or other organizational structure. The term "payer" will be used to refer to an organization that pays one or more providers for the services they deliver. When a patient pays directly for all or most of the cost of a healthcare service, either because they have no health insurance, because the service is not covered by health insurance, or because they have to meet a deductible before their health insurance pays for services, then the patient may also be the payer.
2. These measures are frequently described as the "cost of care," but this is a misnomer because they typically do not measure the actual *cost* of delivering the services, but rather what is *paid* for those services by payers. This represents the cost to the payer, but not the cost to the provider. The distinction is important, because in some cases, payers may be paying less for a service than it actually costs the provider to deliver it, which in many cases forces providers to charge more for other services (or to charge more for the same services to other payers) in order to cover these losses, and in some cases, low payment may result in a shortage of those services. In other cases, providers may be charging significantly more for services than it costs simply because they can.
3. More information about the Medicare QRUR program is available at http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/PhysicianFeedbackProgram/.
4. The Hospital Compare website can be accessed at http://www.medicare.gov/hospitalcompare.
5. Centers for Medicare and Medicaid Services. Measure information form collected for: CMS efficiency measures (claims based). Available at http://www.qualitynet.org.
6. Information on the Medicare Value-Based Payment Modifier for physicians is available at http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/PhysicianFeedbackProgram/index.html. Information on the Medicare Hospital Value-Based Purchasing Program is available at http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/hospital-value-based-purchasing/index.html.
7. Information on the Comprehensive Primary Care Initiative is available at http://innovation.cms.gov/initiatives/Comprehensive-Primary-Care-Initiative/.
8. Centers for Medicare and Medicaid Services. Medicare shared savings program: Accountable care organizations. Federal Register 76 (212): 67802-67990.
9. Information on the Pioneer ACO Model is available at http://innovation.cms.gov/initiatives/Pioneer-ACO-Model/.
10. Narrow networks can also be created based on which providers are willing to charge lower prices or give bigger discounts to the payer; in effect, these price concessions allow providers to modify the spending in the spending measure so that they will rank more favorably relatively to other providers, and thereby be eligible to deliver services to more of the payer's patients.
11. In various circumstances, medical and other healthcare services can be delivered by nurse practitioners, physician assistants, nurse anesthetists, etc. in addition to physicians. Laws vary from state to state as to the extent to which these individuals can order or deliver services without supervision by a physician. For simplicity, the term "physician" will be used in this report to describe healthcare practitioners who independently deliver or order healthcare services.
12. Many types of preventive healthcare services may have little impact on spending today, but can prevent significant expenditures in the future.
13. A study done for the Medicare Payment Advisory Commission found that in six different metropolitan areas studied, 12-22% of total spending for Medicare beneficiaries was associated with episodes in which ten or more physicians were involved, and the spending per episode in these cases averaged between $8,500 and $11,000, compared to an average of a few hundred dollars for the cases when only one physician was involved. Houchens RL, McCracken S, Marder W, Kelley R, Anderson S. Multiple attribution of episodes for physician profiling in Medicare: A preliminary investigation. Medicare Payment Advisory Commission. June 2009.
14. From a payer's perspective, if all of the physicians and other providers who deliver healthcare services to the patient are part of a single healthcare organization, accountability could simply be assigned to that organization. However, that does not resolve the question of how the healthcare organization will assign accountability to the different providers who are involved; it simply transfers that challenge from the payer to the healthcare organization. Nor does it help a patient determine which physician within the healthcare organization is likely to deliver the care they need at a lower cost.
15. "Narrow network" health plans may limit the universe of providers a patient may see, but they do not typically place limits on how many primary care physicians or other physicians the patient may see within the network.

16. CMS defines primary care services using certain procedure and revenue codes. Federal Register 76 (212): 67975.

17. CMS defines this as physicians with a specialty of internal medicine, general practice, family practice, or geriatric medicine. Federal Register 76 (212): 67975.

18. The physician who delivers the "plurality" of services delivers more than any other physician, even if that physician delivers less than a majority of all services the patient receives.

19. Centers for Medicare and Medicaid Services. Medicare shared savings program: Accountable care organizations. Federal Register 76 (212): 67802-67990.

20. Mehrotra A, Adams J, Thomas JW, McGlynn EA. The effect of different attribution rules on individual physician cost profiles. Ann Intern Med. 2010; 152(10):649-654. Four attribution rules were tested based on whether *a single provider* billed the (1) majority of costs, (2) plurality of costs, (3) majority of visits, or (4) plurality of visits for the patient, and two rules were tested based on whether *any provider* billed (5) more than 30% of costs or (6) more than 30% of visits.

21. Chronic disease episodes are typically defined in 12 month periods because patients are generally permitted to change insurance coverage annually.

22. Commonly used groupers are the Symmetry Episode Treatment Groups (ETG) from OptumInsight (https://etg.optum.com/etg-links/episode-treatment-groups/) and Medical Episode Groups (MEG) from Truven Health Analytics (http://truvenhealth.com/your-healthcare-focus/government/medical-episode-grouper-government).

23. For example, one grouper reported 233,673 episodes of coronary artery disease with average spending per episode of $3,998, while the other reported only 201,936 episodes of coronary artery disease with average spending of $3,079 for the same population of patients. Conversely, the first grouper reported only 68,704 episodes of bacterial pneumonia with average spending per episode of $3,054, while the second reported 74,890 episodes of bacterial pneumonia with average spending per episode of $4,427. Medicare Payment Advisory Commission. June 2006 Report to the Congress: Increasing the value of Medicare. Chapter 1: Using episode groupers to assess physician resource use.

24. MaCurdy T, Kerwin J, 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.

25. Rosen A, Liebman E, Aizcorbe A, Cutler DM. Comparing commercial systems for characterizing episodes of care. U.S. Bureau of Economic Analysis Working Paper Series. June 2012.

26. Mehrotra A, *et al., op cit.*

27. An alternative approach that is now being used in conjunction with "shared savings" payment incentives is to determine whether spending attributed to a provider increased by less than spending increased for a comparison group of patients on the same measure (or equivalently, but more rarely, if spending decreased by more than the decrease in spending for the comparison group of patients).

28. Information on the CMS HCC risk adjustment system is available at http://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Risk-Adjustors.html.

29. Schone E, Brown RS. Risk adjustment: What is the state of the art, and how can it be improved? Robert Wood Johnson Foundation. July 2013.

30. Centers for Medicare and Medicaid Services. CMS standardization methodology for allowed amount. Available at http://www.qualitynet.org.

31. With support from the Robert Wood Johnson Foundation, the Network for Regional Healthcare Improvement (NRHI) and five of its member Regional Health Improvement Collaboratives are implementing a Healthcare Regional Cost Measurement and Transparency Pilot. The goals of this initiative are to: (1) Measure and publicly report the total cost of care and resource use in a standardized way across five regions; (2) Create a process for benchmarking multi-payer commercial healthcare costs; (3) Identify the best ways to share cost information with key stakeholders in local communities to identify drivers and reduce healthcare costs; and (4) Conduct focused work with physicians to help them use cost information to adopt practices that will reduce costs and improve care, and to encourage them to serve as leaders in their communities. More information on the initiative is available from NRHI at www.NRHI.org.

32. Pham H, Schrag D, O'Malley AS, Wu B, Bach PB. Care patterns in Medicare and their implications for pay for performance. N Engl J Med 356;11 p. 1130. March 15, 2007.

33. Mehrotra A, *et al., op cit.*

34. HealthPartners. Attribution use in total cost of care: An observational study of commercial administrative methods. 2014.

35. Mehrotra A, *et al., op cit.*

36. Medicare Payment Advisory Commission. June 2006 Report to the Congress. *op cit*.

37. HealthPartners, *op cit.*

38. Houchens RL, *et al. op cit*.

39. The same problem occurs with quality measures based on claims data. If patient health status is only measured for patients who make a visit to the physician's office, a physician who does a better job of keeping patients well will have a higher share of office visits devoted to patients who are sick, making it appear that the physician is actually doing a worse job of improving her patient's health.

40. McWilliams JM, Chernew ME, Dalton JB, Landon BE. Outpatient care patterns and organizational accountability in Medicare. JAMA Intern Med. April 21, 2014.

41. Yalowich R, Wirth B, Takach M. Matching patients with their providers: Lessons on attribution and enrollment from four multi-payer patient-centered medical home initiatives. National Academy for State Health Policy. May 2014. Available at: http://www.nashp.org/sites/default/files/PCMH_Attribution_and_Enrollment.pdf.

42. Mehrotra A, *et al., op cit.*

43. McWilliams JM, *et al., op cit*.

44. CMS stated that it believes the physician group providing the plurality of care during a hospitalization "will be best able to coordinate care and discharge and reduce fragmentation and unnecessary service provision." Centers for Medicare and Medicaid Services. Medicare program; revisions to payment policies under the physician fee schedule, clinical laboratory fee schedule & other revisions to Part B for CY 2014; final rule. 78 Federal Register 74778.

45. Mehrotra A, *et al., op cit.*

46. As the Medicare Payment Advisory Commission stated in one of its reports, "Our analysis found that the key factor in attributing episodes to physicians is the threshold for attribution…further discussion is needed on the extent to

which accountability should be derived from these types of accountability rules. In a payment system as fragmented as Medicare FFS, a physician with 30 percent of the E&M visits in a given episode may not necessarily be aware of the kind of care being provided in the other E&M visits." Medicare Payment Advisory Commission. June 2006 Report to the Congress, *op cit*.

47. Although post-acute care that follows a hospitalization should logically be connected to the hospitalization, one cannot assume that episode grouper software will, in fact, connect them. The MaCurdy study cited earlier found that for Medicare beneficiaries, the Symmetry ETG and Medstat MEG groupers only linked skilled nursing facility stays to the hospital stays that preceded them in half of the episodes analyzed, even though the Medicare program requires an immediately preceding hospital stay in order to pay for a stay in a skilled nursing facility.

48. It is important to note that this discussion relates to assigning accountability for episodes to providers *retrospectively* without any effort to *prospectively* change the way they are paid. If the surgeon had agreed in advance to accept an episode payment with accountability for all of the costs of the episode, that surgeon would then have been able to have taken steps in advance to control the costs that occurred within the episode. See Section VI for a discussion of more accountable payment methods.

49. The newest QRUR reports from Medicare provide significantly more actionable information, including breaking down spending into episodes and then further breaking down episode spending into services delivered or ordered by the provider who is attributed the episode versus services delivered or ordered by others. However, many commercial payers are only producing simple reports comparable to what is shown in Figure 5. More information about the Medicare QRUR program is available at http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/PhysicianFeedbackProgram/

50. Cassel CK, Jain SH. Assessing individual physician performance: Does measurement suppress motivation? JAMA 307(24) 2595-2596.

51. Pham, *et al., op cit.*

52. Mehrotra A, *et al., op cit.*

53. McWilliams *et al., op. cit.*

54. See, for example, Carreyou J, Stewart CS, Barry R. Taxpayers face big Medicare tab for unusual doctor billings. Wall Street Journal, June 9, 2014.

55. Author's calculations from the 2012 National Health Expenditure Accounts. Available from: http://www.cms.gov/NationalHealthExpendData/02_NationalHealthAccountsHistorical.asp#TopOfPage

56. See the information supplied through the Choosing Wisely® campaign at http://www.abimfoundation.org/Initiatives/Choosing-Wisely.aspx.

57. Although most reporting and payment systems that measure spending also attempt to use measures of quality to avoid stinting, there are not quality measures available for all aspects of care, and there would be significant costs associated with attempting to measure all types of quality individually.

58. Since patients should get a colonoscopy every 10 years, this means that 10% of the patients would receive a colonoscopy every year, assuming the patients' previous colonoscopies were distributed evenly across previous years.

59. In the Medicare Physician Value-Based Modifier program, the Centers for Medicare and Medicaid Services classifies physicians as "high" or "low" based on whether measures of spending and quality are at least one standard deviation above or below the mean for all physicians being measured. The example in the text is over-simplified because it only considers primary care visits and colonoscopies as components of total spending, but the same phenomenon could occur if a broader range of services and spending were considered.

60. Information on the CMS HCC risk adjustment system is available at http://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Risk-Adjustors.html.

61. Prospective risk adjustment may also *overestimate* spending in some cases if a health issue the patient had in the previous year was resolved or had been mitigated but that issue was still incorporated into the risk score.

62. As more health plans begin receiving risk adjusted payments through insurance exchanges and Medicare Advantage, accurate and complete diagnosis coding will become more important to them. Some health plans are already encouraging providers to fill out additional claims forms if they need to record additional diagnosis codes.

63. 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_Trnsfrming_Care.pdf.

64. Noyes K, Liu H, Temkin-Greener H. Medicare capitation model, functional status, and multiple comorbidities: model accuracy. Am J Manag Care. 2008 October; 14(10): 679-690.

65. Hibbard JH, Greene J, Overton V. Patients with lower activation associated with higher costs; delivery systems should know their patients' 'scores.' Health Aff. February 2013; 32(2): 216-222.

66. Adams, JL. The reliability of provider profiling: A tutorial. RAND Health, 2009. Available at: http://www.rand.org/pubs/technical_reports/TR653.html.

67. See, for example, Adams JL, Mehrotra A, Thomas JW, McGlynn EA. Physician cost profiling – reliability and risk of misclassification. N Engl J Med 362(11):1014-1021, and Houchens RL. The reliability of physician cost profiling in Medicare. Thomson Reuters. August 2010.

68. Yu H, Mehrotra A, Adams J. Reliability of utilization measures for primary care physician profiling. Healthcare 1 (2013) 22-29.

69. Yu H, *et al., op cit*.

70. HealthPartners. Total cost of care bootstrap reliability analysis. HealthPartners 2014. Available at: https://www.healthpartners.com/ucm/groups/public/@hp/@public/documents/documents/dev_057895.pdf.

71. Of the five providers with the lowest per patient spending in 2007 out of the 19 measured (i.e., the lowest 25%), only two remained in the lowest 25% in 2009, and two had moved to the most expensive half of the providers. The study stated that "the factors that drive variation between years within a provider are cost per unit control and resource management," but no data were presented on the causes of variation. HealthPartners, op cit., page 5.

72. Of the five providers with the lowest resource use per patient in 2007 out of 19 measured, only two remained in the lowest 25% on resource use by 2009, and one had the highest resource use in 2009 among all 19 providers. HealthPartners. Total resource use bootstrap reliability analysis. HealthPartners 2014. Available at: https://www.healthpartners.com/ucm/groups/public/@hp/@public/documents/documents/dev_057896.pdf.

73. MaCurdy T, Shafrin J, Hartmann E, Ho M, Talbot L, Ueda K, Zhang Z. Evaluating the stability of physician efficiency scores. Acumen LLC. February 2010. Available at: http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Reports/downloads/StabilityinPhysicianScores_2010.pdf.

74. McWilliams JM *et al, op cit*.

75. In some cases, claims forms may not clearly distinguish which *individual physician delivered* a service, only which *physician practice billed* for the service.

76. It is important to distinguish payments used to pay for drugs from payments for the physician's or staff's time. When the Centers for Medicare and Medicaid Services released data on physician billings in the spring of 2014, it did not make clear that some payments were for a physician's time and others were for drugs used by the physician, leading many people to erroneously conclude that oncologists and ophthalmologists were earning more money than other physicians.

77. Silver J. Drugs for macular degeneration, price discrimination, and Medicare's responsibility not to overpay. JAMA May 23, 2014.

78. Greene RA, Beckman HB, Mahoney T. Beyond the efficiency index: Finding a better way to reduce overuse and increase efficiency in physician care. Health Aff. May 2008; 27(4): w250-w259.

79. Dowd B, Li C, Swenson T, Coulam R, Levy J. Medicare's physician quality reporting system (PQRS): Quality measurement and beneficiary attribution. Medicare and Medicaid Research Review. 2014: 4(2).

80. This does not assume that the actual payment is made on a bundled basis. However, as discussed in Section VI, reporting the data in bundles facilitates movement toward bundled payments.

81. The standard billing form for physicians has a field where physicians identify where they performed the service, but it does not say what that facility was paid for its services. The standard billing form for inpatient and outpatient facilities (the CMS 1450/UB-04) has a field for recording the attending physician, but it does not say what the physician was paid for his or her services. This facility form also has spaces to indicate other physicians who were involved, but commercial health plans may not require that this field be completed.

82. Because episode groupers typically base their groupings on diagnosis codes rather than procedure codes, they do not always combine services in logical bundles. The Acumen study cited earlier found that for hospital stays for Medicare beneficiaries, in more than half of the cases, the Symmetry ETG Grouper and the Medstat MEG grouper failed to link the bills for the visits physicians made during the patient's hospital stay to the hospital bill for the stay. MaCurdy T, Kerwin J, *et al. op cit*.

83. In 2012, 64% of Medicare beneficiaries received an outpatient hospital service but only 18% received an inpatient hospital service. Author's calculations based on Medicare Geographic Variation Public Use Files.

84. In Medicare, the physician performing the procedure and the hospital where an outpatient procedure was performed would generally be expected to bill for the service using the same CPT (Current Procedural Terminology™) code with the same service date, but other physicians or providers involved in the patient's care may use different codes to represent their specific roles in the overall bundle of services delivered.

85. For example, see: Brill JV *et al*. A bundled payment framework for colonoscopy performed for colorectal cancer screening or surveillance. Gastroenterology 2014; 146:849-853; Williams T, Robinson J. Bundled episode-of-care payment for orthopedic surgery: the Integrated Healthcare Association initiative. Integrated Healthcare Association. September 2013; Rastogi A *et al*. Prometheus payment model: application to hip and knee replacement surgery. Clin Orthop Relat Res. June 2009.

86. Most hospitals are paid by Medicare through the Inpatient Prospective Payment System, which provides a single payment for all of the services delivered during the inpatient stay based on the Diagnosis Related Group assigned to the admission.

87. Physicians bill for services using Common Procedural Terminology™ codes, but hospital payments are based on procedure codes and diagnosis codes from the International Classification of Diseases – Clinical Modification.

88. There are differences in the amounts Medicare pays to different providers for the same service; for example, teaching hospitals are paid more than community hospitals in the same community, providers in some geographic regions are paid more than others, etc. However, because these differences are determined by formula, they can be removed in order to compute "price-standardized" spending comparisons across providers. This approach may or may not be adequate to adjust for appropriate differences in the costs of these different providers, as discussed in more detail in Section IV.

89. This is not intended to imply that Medicare payment levels are, in fact, the appropriate payment levels; some of them are believed to be too high and others are believed to be too low. However, in the absence of more direct information about what it actually costs to deliver a service, imperfect proxies may have to be used.

90. For example, the Total Care Relative Resource Value™ measure developed by HealthPartners and endorsed by the National Quality Forum calculates the value of individual services based in part on the relative weights among services used by Medicare, but not on the actual Medicare payment levels. HealthPartners. HealthPartners white paper: Total Care Relative Resource Value (TCRRV). Available at: http://www.healthpartners.com/tcoc.

91. The cost of delivering services is not independent of volume, so one cannot assume that if fewer services are ordered, the unit cost of the remaining services will stay unchanged or that the current price per unit will be adequate to cover costs when volumes decline. Many providers have significant fixed costs, and if the volume of services ordered decreases, the unit costs of the remaining services will increase, so the provider may need to raise their prices (if they have the ability to do so) in order to generate sufficient revenue to continue covering the fixed costs.

92. Medicare pays a gastroenterologist more for a procedure done in the office because the payment is designed to cover the extra costs the gastroenterologist incurs to equip and staff their office in order to do procedures there.

93. This is no different than in other industries, where people buy a product based on whether they think the price is justified by the usefulness of the product, not based on how much the personnel producing the product are paid.

94. Centers for Medicare and Medicaid Services. Medicare claims processing manual. Chapter 26 – Completing and processing form CMS-1500 data set. Available at: http://www.cms.gov/Regulations-and-Guidance/Guidance/Manuals/downloads/clm104c26.pdf.

95. Centers for Medicare and Medicaid Services. Detailed methods of the 2012 medical group practice supplemental quality and resource use reports (QRURs). June 2014. Available at http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/PhysicianFeedbackProgram/Downloads/2012-Supplemental-QRURs-Methods.pdf .

96. National Uniform Claim Committee. 1500 health insurance claim form reference instruction manual. Available at http://www.nucc.org/images/stories/PDF/claim_form_manual_v9-0_7-13.pdf.

97. It is, of course, quite possible that a physician would order a test prior to seeing the patient for an evaluation and management visit, so a variant of the rule would be to include laboratory testing that occurs shortly before an E&M visit. Since no rule will be perfect, it would be preferable to have information on the ordering provider accurately recorded on the claim.

98. This is consistent with the way that CMS is using episode measures in the Hospital Value-Based Purchasing program. If a patient is admitted to a different hospital for a complication associated with a previous hospitalization, both the original hospitalization and the readmission are included in an episode assigned to the initial hospital, but a second episode is also created for the readmission that does not include the original hospitalization, and this second episode is assigned to the hospital where the patient was readmitted. MaCurdy T, Perlroth D, *et al*. Methodology for developing the six hospital-based episode measures: Supplemental documentation for the fiscal year (FY) 2015 Inpatient Prospective Payment System and Long-Term Care Hospital Prospective Payment System Proposed Rule. Available at: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/hospital-value-based-purchasing/Downloads/Hospital-Based-Episode-Measures-Supplemental-Documentation.pdf.

99. This is also consistent with the way that CMS is using episode measures in the Hospital Value-Based Purchasing program, since a readmission is not only included in an episode for the initial hospitalization but is also considered to begin a new episode of care that is assigned to the second hospital.

100. A study done for CMS by Acumen LLC found that outpatient visits following hospitalizations for hip fractures were often defined as new episodes rather than as part of the original hip surgery. The average episode spending was then based on a mix of expensive episodes (the hospitalizations) and inexpensive episodes (the outpatient visits afterward), which in turn made it appear as if the episodes which included hospitalizations for hip fractures were unusually expensive. MaCurdy T, Shafrin J, et al. Challenges in the risk adjustment of episode costs. Acumen, LLC February 2010.

101. For example, the PROMETHEUS payment methodology defined a way of apportioning accountability among multiple providers involved in an episode if the providers themselves do not define a different approach. de Brantes F, Gosfield AG, et al. Sustaining the medical home: How PROMETHEUS Payment can revitalize primary care. PROMETHEUS Payment Inc. and Robert Wood Johnson Foundation. Available at http://www.hci3.org/sites/default/files/files/PROMETHEUS%20-%20Medical%20Home%20-%20full%20packet%20-%20FINAL.pdf.

102. Some recommended services can reduce spending in the long run but increase spending in the short run. Measuring spending on an annual basis can unintentionally discourage spending on preventive services and lead to higher spending and worse patient outcomes in the long run.

103. Information on the Choosing Wisely® campaign is available at http://www.abimfoundation.org/Initiatives/Choosing-Wisely.aspx.

104. 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.

105. The Washington Health Alliance, the Washington State Hospital Association, and the Washington State Medical Association have developed specifications for identifying some of these services through claims data, but they caution that the specifications should only be used for analyses at the regional level and they prohibit using them at the individual medical group level. Washington State Choosing Wisely Task Force. Choosing Wisely claims-based technical specifications. Available at: http://wahealthalliance.org/wp-content/uploads/2013/11/Choosing_Wisely_Specifications_2014.pdf.

106. Goldfield N *et al, op cit.*

107. Hibbard JH, Greene J, Sofaer S, *et al*. An experiment shows that a well-designed report on costs and quality can help consumers choose high-value health care. Health Aff 2012; 31:560-568.

108. de Brantes F, D'Andrea G, Rosenthal MB. Should healthcare come with a warranty? Health Aff 28(4): w678-w687. 2009.

109. 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.

110. Goldfield N. *et al*. *op cit*.

111. This could occur because of patient preference (not wanting to return to the same hospital that caused the problem) or because the patient received the initial treatment at a hospital further from home and was taken to the nearest hospital for diagnosis and treatment of problems that occurred after discharge.

112. For example, the total cost of the avoidable service could be allocated among the physicians based on the proportion of total evaluation and management services they each billed during a defined period of time prior to the avoidable service, and services could be allocated between a hospital and a post-acute care facility based on the number of days spent in each facility prior to the condition occurring. The Potentially Avoidable Complications (PAC) definitions developed by the Health Care Incentives

Improvement Institute (HCI3) have two features that make them particularly useful for apportioning accountability among different providers. First, different sets of PACs are associated with specific procedures or patient conditions, making it easier to assign them to different physicians (for example, a hospitalization for an exacerbation of COPD is defined as a PAC for a patient with COPD, but not for a patient with diabetes, so the former could be assigned to a pulmonologist, but not the latter). Second, PACs are further disaggregated into three subcategories: (1) PACs associated with the "anchor condition," (2) PACs associated with comorbidities, and (3) PACs associated with patient safety failures. For patients with multiple conditions and patients being treated by multiple providers, this improves the ability to associate PACs with the providers who would be most likely to have caused them or most likely to have been able to prevent them.

113. MaCurdy T, Perlroth D, *et al. op cit*.

114. Information on the Inpatient Prospective Payment System is available at http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS/IPPS-Regulations-and-Notices.html.

115. In its regulations for the Inpatient Prospective Payment System, CMS has 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.

116. The Johns Hopkins ACG System: Technical Reference Guide. Johns Hopkins Bloomberg School of Public Health. Available at: http://acg.jhsph.org/public-docs/ACGv10.0TechRefGuide.pdf.

117. Hughes JS, Averill RF, Eisenhandler J, Goldfield NI, *et al*. Clinical risk groups (CRGs): a classification system for risk-adjusted capitation-based payment and health care management. Medical Care 42(1): pp. 81-90. January 2004.

118. The fact that a large number of categories can be defined does not mean they all need to be used in every application. However, the larger the number of categories, the greater the ability to define patient groups that are similar on the characteristics most relevant for particular types of specialists or procedures. For example, the detailed categories in the CRG model allow spending to be examined separately for patients who are undergoing dialysis, who have quadriplegia or paraplegia, or who have metastatic malignancies if those characteristics are relevant to the services being provided.

119. Hughes JS, *et al. op cit.*

*120. Ibid*. In the Johns Hopkins ACG system, the ACG categories are grouped into 6 different "Resource Utilization Bands" based on the expected relative resource use of patients. The ACG categories can be subdivided into 23 different MACs, 32 different ADGs, and 12 "Collapsed" ADGs.

121. For example, in a comparison of different risk adjustment models conducted by the Society of Actuaries, many models did more poorly in predicting spending for patients with breast cancer, heart disease, HIV, and mental illness than for patients with asthma or diabetes. Winkelman R, Mehmud S. A comparative analysis of claims-based tools for health risk assessment. Society of Actuaries, April 20, 2007.

122. Winkelman R, Mehmud S. *op cit.*

123. For example, if $250,000 were actually spent on a patient with characteristics that typically would result in spending of, say, $10,000, only $100,000 for that patient might be factored into the calculation of average spending per patient. The truncation limit is often set based on the distribution of actual spending for all similar patients, e.g., the truncation amount might be set at the 99th percentile.

124. Under this approach, there will still be outlier patients, i.e., patients who have unique or rare combinations of characteristics, but there will likely be fewer of them than under typical risk adjustment approaches, and it will be clearer that the appropriate approach is to exclude them from spending comparisons.

125. The example assumes that the patient has a rare but identifiable condition. In many cases, a patient might require additional services due to factors that cannot clearly be identified. The approach described would allow patients with identifiable but rare conditions to be analyzed separately, so that there would be a smaller number of outlier cases.

126. Different weighting schemes can create different types of incentives for providers to focus care on different kinds of patients. See, for example, MaCurdy T, Shafrin J, Zheng D. Optimal pay-for-performance scores: How to incentivize physicians to behave efficiently using episode-based measures. Acumen, LLC February 2011. Available at: https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Reports/downloads/MaCurdy_Incentivize_Physicians_Optimal_P4P_Scores_Feb_2011.pdf.

127. In contrast, applying the same average risk-adjustment formula to every patient would mask the fact that one provider has completely different types of patients than another, or that providers have very small numbers of patients with certain characteristics.

128. Winkelman R, Mehmud S. *op cit.*

129. In most fee for service systems, a physician's payment is based on what procedure they performed, not on the precise diagnoses of the patient's conditions. There is typically no reason for physicians to record all of the diagnoses for a patient in claims for payment, particularly if the specific service the physician provided was only related to one of those diagnoses. In contrast, the Diagnosis Related Group system that is used by Medicare and many commercial payers to pay hospitals for inpatient stays explicitly bases payment in part on the diagnoses reported for a patient.

130. MaCurdy T, Perlroth D, *et al., op. cit.* Centers for Medicare and Medicaid Services. Detailed methodology for the 2012 quality and resource use reports. September 2013. Available at: http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/PhysicianFeedbackProgram/Downloads/PY2012-Detailed-Methods.pdf.

131. The weakness in this approach is that it may use too narrow a window to identify diagnoses, if there is information about the patient's condition that is only reflected on

claims filed more than six months before the episode began.

132. As noted in Section IV-A, in addition to clarifying diagnostic information, clinical information from EHRs can also help to identify patients who are being cared for by a provider but not receiving billable services.

133. This does not mean that the entire differential between a teaching hospital's costs and the costs of other providers would be adjusted away, only the differential that was believed to be due directly to the teaching and research functions. A common approach is to use the explicit payment differentials that Medicare has defined for teaching hospitals, providers located in high cost areas, etc. as proxies for these differences.

134. The measure is valid as a measure of how much is spent on the *patient*, but not as a measure of how much spending an individual *provider* could have controlled.

135. There are mechanisms today to enable physicians and other providers to correct errors in claims data used to generate quality metrics. For example, the federal regulations governing Medicare Qualified Entities require them to "confidentially share measures, measurement methodologies, and measure results with providers and suppliers at least 60 calendar days before making reports public" and to "allow providers and suppliers the opportunity to make … a request for error correction." (76 FR 235 p. 76570) The same mechanisms could be used to enable correction of errors in claims data used to generate spending breakdowns.

136. The report in Figure 21 shows differences in spending in percentage terms and absolute dollar amounts, since each of these independently indicate the magnitude of the opportunity for achieving savings. If a category has a higher percentage difference, it suggests that there may be easier opportunities to reduce spending in that category. However, if the absolute amount of spending in that category is low, then even a big reduction in the percentage difference may result in a relatively small dollar savings.

137. A more detailed discussion of this process for developing a business case for payment reform is available in Miller HD. Making the business case for payment and delivery reform. Robert Wood Johnson Foundation and Network for Regional Healthcare Improvement. February 2014. Available from http://www.chqpr.org/reports.html.

138. Many providers have significant fixed costs, and if the volume of services delivered decreases, the unit costs of the remaining services will increase, particularly in the short run.

139. For more detail on different types of payment systems, see Miller HD. Transitioning to accountable care: Incremental payment reforms to support higher quality, more affordable healthcare. Pittsburgh, PA: Center for Healthcare Quality and Payment Reform; 2012. Available from: http://www.chqpr.org/reports.html.

140. For more information on the ACE Demonstration, see http://www.cms.gov/Medicare/Demonstration-Projects/DemoProjectsEvalRpts/Medicare-Demonstrations-Items/CMS1204388.html.

141. For episode payments, the Integrated Healthcare Association and the Health Care Incentives Improvement Institute have both developed detailed specifications as to which services will be included and which will not, as well as risk adjustments, risk exclusions, etc. For more information, see http://www.iha.org/bundled-payment-implementation.html and http://www.hci3.org/what_is_prometheus

142. See, for example, Paulus RA, Davis K, Steele GD. Continuous innovation in health care: Implications of the Geisinger experience. Health Aff 27(5): 1235-1245. September/October 2008.

143. For more information on the CMS Bundled Payments Initiative, see http://www.innovations.cms.gov/initiatives/Bundled-Payments/index.html

144. More information on Regional Health Improvement Collaboratives can be obtained from the Network for Regional Healthcare Improvement (http://www.NRHI.org ).

145. More information on the Qualified Entity program is available at http://www.cms.gov/Research-Statistics-Data-and-Systems/Monitoring-Programs/QEMedicareData/index.html?redirect=/QEMedicareData.

146. This is generally referred to as the "allowed amount" for a service, in contrast to the higher amount that the provider "charged" for a service.

147. Beckman HB. Lost in translation: Physicians' struggle with cost-reduction programs. Ann Int Med 154(6) 430-433. Also, Cassel CK, Jain SH. *op cit.*

## ACKNOWLEDGEMENTS

Many of the issues discussed in this report were identified through the author's involvement over several years in payment reform projects supported by the Robert Wood Johnson Foundation and in communities working on healthcare payment and delivery reform as part of the RWJF-sponsored Aligning Forces for Quality Program. The opportunities to participate in these projects and the ongoing encouragement, advice, and support of the staff at the Robert Wood Johnson Foundation, particularly Andrea Ducas, Michael Painter, and Anne Weiss, are greatly appreciated.

The report benefited from helpful comments on an earlier draft provided by Howard Beckman, Michael DeLorenzo, Andrea Ducas, Richard Fuller, Norbert Goldfield, Sandy Marks, Karen Nychay, Louise Probst, Karen Timberlake, Meredith Roberts Tomasi, and Dolores Yanagihara. The author is solely responsible for any errors, however.

## ABOUT THE AUTHOR

Harold D. Miller is the President and CEO of the Center for Healthcare Quality and Payment Reform. Miller also serves as Adjunct Professor of Public Policy and Management at Carnegie Mellon University.

Miller is a nationally-recognized expert on healthcare payment and delivery reform, and has given invited testimony to Congress on how to reform healthcare payment. He has authored a number of papers and reports on health care payment and delivery reform, including "From Volume to Value: Better Ways to Pay for Healthcare," which appeared in the September 2009 issue of *Health Affairs*, "Win-Win-Win Approaches to Healthcare Cost Control Through Physician-Led Payment Reform, which appeared in the March 2014 issue of *Clinical Gastroenterology and Hepatology*, the Center for Healthcare Quality and Payment Reform's reports *How to Create Accountable Care Organizations, Transitioning to Accountable Care*, and *Ten Barriers to Healthcare Payment Reform and How to Overcome Them*, the Network for Regional Healthcare Improvement's report *Making the Business Case for Payment and Delivery Reform*, the Massachusetts Hospital Association's report *Creating Accountable Care Organizations in Massachusetts*, the American Medical Association's report *Pathways for Physician Success Under Healthcare Payment and Delivery Reforms,* the Medical Society of Virginia's report *How Virginia Physicians Can Improve the Quality and Reduce the Costs of Health Care Through Payment and Delivery Reforms,* and the Association of Departments of Family Medicine's report *Leading the Way in Accountable Care: How Departments of Family Medicine Can Help Create a Higher Quality, More Affordable Healthcare System.*

From 2008 to 2013, Miller served as the President and CEO of the Network for Regional Healthcare Improvement (NRHI), the national association of Regional Health Improvement Collaboratives, and he organized NRHI's national Summits on Healthcare Payment Reform in 2007 and 2008 and its Summit on Regional Healthcare Transformation in 2013. His report *Creating Payment Systems to Accelerate Value-Driven Health Care: Issues and Options for Policy Reform* which was prepared for the 2007 Summit was published by the Commonwealth Fund in September, 2007, and his summary of the recommendations from the 2008 Payment Reform Summit, *From Volume to Value: Transforming Healthcare Payment and Delivery Systems to Improve Quality and Reduce Costs*, was published in November 2008 by NRHI and the Robert Wood Johnson Foundation.

Miller also served for several years as the Strategic Initiatives Consultant to the Pittsburgh Regional Health Initiative (PRHI). His work demonstrating the significant financial penalties that hospitals can face if they reduce hospital-acquired infections was featured in *Modern Healthcare* magazine in December, 2007. He designed and led a multi-year PRHI initiative that significantly reduced preventable hospital admissions and readmissions through improved care for chronic disease patients.

Miller has worked in more than 30 states and metropolitan regions to help physicians, hospitals, employers, health plans, and government agencies design and implement payment and delivery system reforms, and he assisted the Centers for Medicare and Medicaid Services with the implementation of its Comprehensive Primary Care Initiative in 2012. He has served on the Board of Directors of the National Quality Forum since 2009.

MEASURING AND ASSIGNING ACCOUNTABILITY FOR HEALTHCARE SPENDING
First Edition, August 2014

www.CHQPR.org


320 Ft. Duquesne Blvd.,
Suite 20-J
Pittsburgh, PA 15222

VOICE: (412) 803-3650
FAX: (412) 803-3651
EMAIL: Info@CHQPR.ORG
www.CHQPR.org