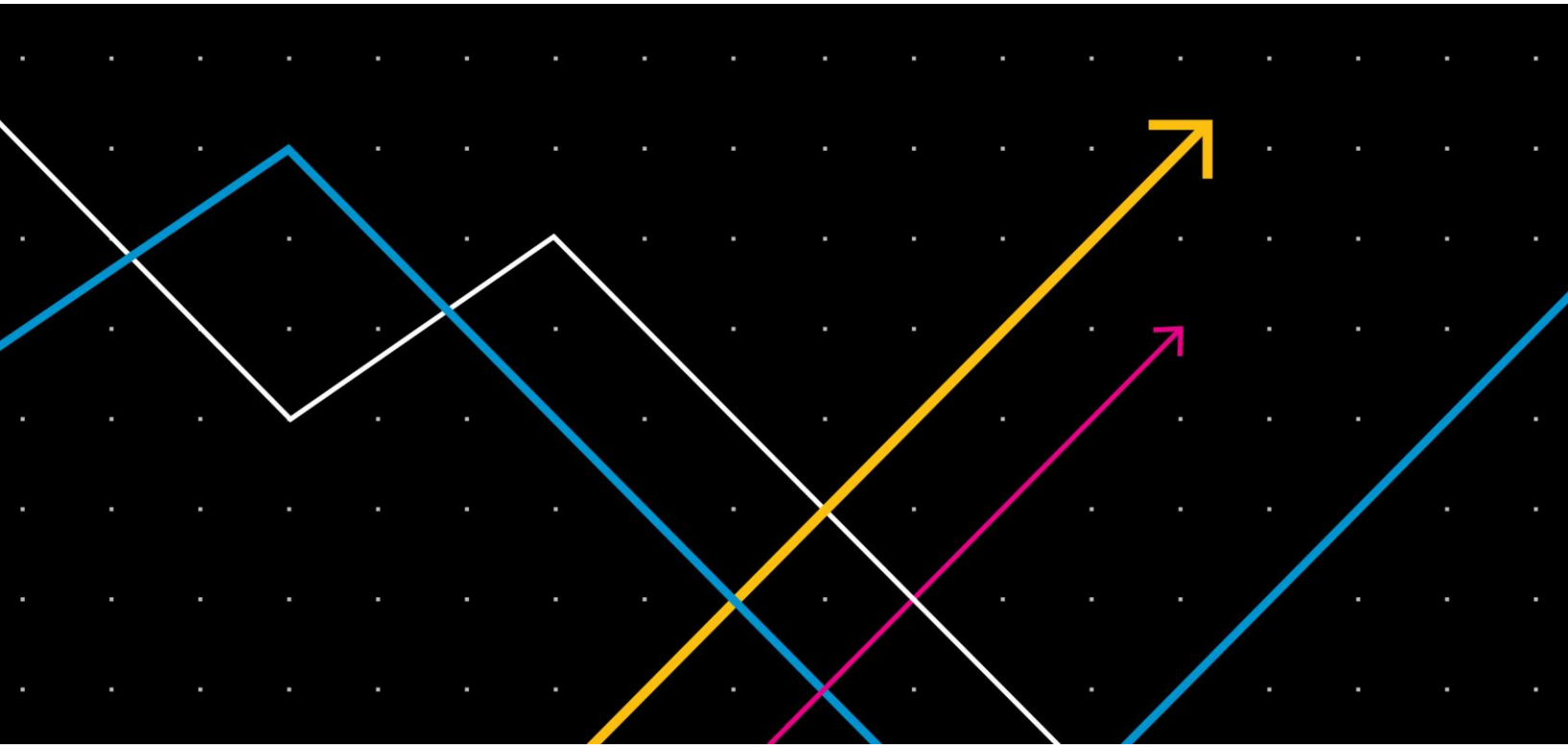

RESEARCH REPORT

Payment Methods and Benefit Designs: How They Work and How They Work Together to Improve Health Care

Matching Payment Methods with Benefit Designs to Support Delivery Reforms

Robert A. Berenson
URBAN INSTITUTE

Suzanne F. Delbanco
CATALYST FOR PAYMENT REFORM

Roslyn Murray
CATALYST FOR PAYMENT REFORM

Divvy K. Upadhyay
URBAN INSTITUTE

April 2016

URBAN INSTITUTE

OUR HEALTH CARE SYSTEM NEEDS CPR

CATALYST FOR PAYMENT REFORM

ABOUT THE URBAN INSTITUTE

The nonprofit Urban Institute is dedicated to elevating the debate on social and economic policy. For nearly five decades, Urban scholars have conducted research and offered evidence-based solutions that improve lives and strengthen communities across a rapidly urbanizing world. Their objective research helps expand opportunities for all, reduce hardship among the most vulnerable, and strengthen the effectiveness of the public sector.

Contents

Acknowledgments

This report was funded by the Robert Wood Johnson Foundation. We are grateful to them and to all our funders, who make it possible for Urban to advance its mission.

The views expressed are those of the authors and should not be attributed to the Urban Institute, its trustees, or its funders. Funders do not determine research findings or the insights and recommendations of Urban experts. Further information on the Urban Institute's funding principles is available at www.urban.org/support.

A technical expert panel advised the project team and reviewed the reports at different stages. This team consists of **Michael E. Chernew,** Leonard D. Schaeffer professor of health care policy and director of Healthcare Markets and Regulation Lab, Harvard Medical School; **Francois de Brantes,** executive director, Health Care Incentives Improvement Institute; **Anna Fallieras,** program leader, Health Care Initiatives and Policy, General Electric; **Kate Farley,** executive director, Pennsylvania Employees Benefit Trust Fund; **Joseph J. Fifer,** president and chief executive officer, Healthcare Financial Management Association; **Robert Galvin,** chief executive officer, Equity Healthcare, operating partner, Blackstone, and former chief medical officer, General Electric; **Paul Ginsburg,** professor and director of public policy, Schaeffer Center for Health Policy and Economics, University of Southern California, and senior fellow and director, Center for Health Policy, Brookings Institution; **Stuart Guterman,** senior scholar in residence, AcademyHealth, and former vice president, Medicare and Cost Control, The Commonwealth Fund; **Vincent E. Kerr,** president, Care Solutions, National Accounts, UnitedHealthcare, and former chief medical officer, Ford Motor Company; **Peter Kongstvedt,** principal, P.R. Kongstvedt Company, LLC, and senior health policy faculty member, George Mason University; **Jeff Levin-Scherz,** assistant professor, Department of Health Policy and Management, Harvard University, and national coleader, Willis Towers Watson; **Robert Murray,** president and consultant, Global Health Payment LLC, and former executive director, Maryland Health Services Cost Review Commission; **Dave Prugh,** independent adviser and consultant, and former vice president of Reimbursement and Contracting Strategy, WellPoint, Inc.; **Simeon Schwartz,** founding president and chief executive officer, WESTMED Medical Group; and **Lisa Woods,** senior director, US health care, Walmart Stores Inc.

Payment reform promises to substitute value for volume. Yet, value- and volume-based approaches typically are implemented together. All payment methods have strengths and weaknesses, and how they affect the behavior of health care providers depends on their operational design features and, crucially, on how they interact with benefit design. Those seeking greater value for their health care dollar are also turning to innovation in benefit design, which also typically involves the implementation of more than one approach at a time—each with its own strengths, weaknesses, and effect on consumer health care behavior. Although payment and benefit design each has received significant attention independently, the intersection between the two has received little if any. The Urban Institute partnered with Catalyst for Payment Reform to explore how established and proposed payment methods and benefit design options work on their own and together. We also examined how payment and benefit design can be blended to improve health care delivery. All reports and chapters can be found on our project page: Payment Methods and Benefit Designs: How They Work and How They Work Together to Improve Health Care.

The Patient Centered Medical Home—Advanced Primary Care

The patient-centered medical home (PCMH) is a redesign of primary care delivery emphasizing population health management, multidisciplinary teams, and care management for at-risk patients. Many PCMHs today receive payment according to a base of standard fee schedules, along with incremental payments for care coordination. However, a hybrid approach, which demotes the incentive to overgenerate face-to-face visits inherent in fee-for-service and promotes beneficial activities that previously did not qualify for payment (e.g., care coordination, robust e-mail and phone communication), may be more effective. Two examples of payment approaches that might best support a PCMH include (1) a base method that includes a hybrid of a reduced price fee schedule and capitation, with additional incremental payments, such as payment for performance (P4P) and shared savings and (2) primary care capitation as the base method with some incremental payments using P4P and shared savings methods.

While we can change supply-side incentives to focus less on generating visits and more on coordinating care, we can change incentives for consumers to use providers in these primary care models. Physicians are also less willing to take on risk if they lack control over patients' use of health

care services. Therefore, a narrow network design could be beneficial, as it essentially restricts consumers' access to providers to those in-network and to whom they are referred. Second, for PCMHs with capitated base payments, moderate cost-sharing, usually co-payments, for consumers could temper the likelihood to seek health care services they do not need. Value-based insurance design can align both consumers' and providers' interests in primary preventive care and in care outcomes. Capitated payments can also support the use of alternative sites of care—such as retail clinics and telehealth—in the PCMH. Additionally, other utilization management approaches can support providers subject to risk-based payment arrangements in their management of patients and patients' pursuit of health care services.

Introduction

The PCMH is a redesign of primary care delivery, which emphasizes population health management, multidisciplinary teams, and care management for patients at risk of frequent hospitalizations. Many, including the primary care specialties that originated the current medical home movement, consider supportive payment reform as a requirement intrinsic to advancing the concept and implementation of the medical home. The Medicare Access and CHIP Reauthorization Act (MACRA) of 2015 identifies alternative payment models that would provide additional payment to physicians who qualify as patient-centered medical homes. However, consumers need incentives to use these primary care models and help them succeed; certain benefit designs can encourage use.

The Basic Payment Approach

PCMHs are presumed to be one of the approved alternative payment models in MACRA; this creates a dilemma, because the medical home is in fact a model for how to deliver care, not a payment method. PCMHs are supported by a variety of payment methods, and there is active debate about how best to do this. In PCMH demonstrations, payers have adopted a mix of payment methods, typically built on a base of standard fee schedules.

For example, in the eight-state Multi-Payer Advanced Primary Care Demonstration, the common central approach was to pay standard fee schedule payments. All states then provided the medical home with an incremental per capita payment for care management activities for each patient (with the payment variously adjusted for age, number of chronic conditions, the level of medical home

achievement, or another factor). However, each state's approach then varied: some had P4P bonuses related to performance on quality metrics, and one used a shared savings approach related to total cost of care analysis.

Complementary Payment Approaches

As noted above, most payment approaches to the medical home start with standard fee schedules, which are typically modeled closely on Medicare's. Fee schedules have some positive attributes, including the ability to incentivize performance of specific targeted activities and to price items such as vaccines in a way that accounts for price fluctuations. However, fee schedules are the prototypical volume-based payment method, rewarding additional services (needed or not), interfering with efforts to reward greater attention to care coordination across providers and other social service supports, and shortchanging activities—central to the PCMH concept—that cannot be readily recognized and paid on a fee-for-service basis, such as frequent e-mail and telephone exchanges between practices and patients and their caregivers.

Some think the unimpressive results from PCMH pilots to date stem from base and incremental payments being insufficient to permit clinicians to reengineer care delivery, providing primary care services while emphasizing access, continuity, comprehensiveness, and coordination. In addition, the current approaches to layering small, incremental payments on top of a standard fee schedule do not alter primary care practices' incentives to generate face-to-face visits. Accordingly, there is interest in payment methods that move away from a base of standard fee schedule payments, so that physician practices are less constrained in how they deploy resources to achieve the PCMHs' promise to improve patient care.

Two approaches with great potential to support PCMHs include (1) a base method that incorporates a hybrid mix of a reduced fee schedule and capitation, with some additional incremental payments, such as P4P and shared savings and (2) primary care capitation as the base method with some incremental payments using P4P and shared savings.

The first method, reducing the fee schedule payment amounts, could move the incentives in the PCMH to greater neutrality, such that physicians would no longer see face-to-face visits as their only major source of revenue. Capitation—a per member per month, perhaps case-mix-adjusted payment—would support activities that medical homes seek to undertake but that current fee schedules typically do not recognize. Nearly 25 percent of activities undertaken by primary care practices are not covered

by fee schedule payments (Chen et al. 2011; Gilchrist et al. 2005). Presumably, under a full or partial capitation payment approach, even a greater percentage of work activity would reflect activities unrelated to visits.

A modified approach to maintaining fee schedule payments would be to include generous payments for activities that need emphasis, such as immunizations and other preventive services; these are the same services typically subject to external measurement and reward or penalty. Another approach could be to support chronic care management for certain patients through a monthly payment in addition to the monthly capitation payment.

The second method, primary care capitation, was a dominant method in 1980s and 1990s health maintenance organizations (HMOs). The payment method was supported by benefit designs that required members to select a physician practice for routine care and referrals—the so-called gatekeeper. The advantage is that physicians are completely at liberty to decide how to deploy resources—and how to apportion their own time—to best serve the population for whom they are responsible. The concern is that physicians might stint on services or respond to their effective fixed budget by overreferring to other physicians, thereby defeating one core purpose of the medical home. Hence, measuring performance on basic parameters of access and use and possibly adopting incremental payment approaches, such as P4P and shared savings, may be necessary to discipline some potential adverse impacts of pure capitation. Under a predominantly primary care capitation approach, it is still possible to pay fee-for-service for important services, labeled carve-outs or bill-aboves, creating a direct payment incentive for practices.

Of note, compared to the '80s and '90s, we are in a better position to adjust case mixes to determine more accurate capitation rates. We also have better ability to use quality measures to detect providers' stinting on some services, particularly primary and secondary prevention services.

Finally, these alternative payment methods for PCMHs are quite compatible with population-based payment methods, such as shared savings, shared risk, and capitation; they all move away from the volume-based incentives inherent in basic fee schedule payments. For example, it would be conceptually and operationally logical for the intermediary, accountable-care-style organization receiving a globally capitated payment to in turn subcapitate its physicians. Similarly, the organization could adopt the hybrid model to assure the flow of accurate encounter data, which is essential to ACO data monitoring and external oversight.

Complementary Benefits Designs

In general, benefit designs should facilitate ready access for consumers to their medical homes, reducing any financial barriers. The approach will be most successful if the PCMH retains the responsibility for managing referrals, high-cost testing, procedures, and drugs. Therefore, a narrow network would be an appropriate mechanism to limit consumers to seeking care from the PCMH providers and their referrals. With a narrow network approach, patients would not receive coverage for care out of the network, supporting providers' ability to manage their patients' care. Providers would likely be more willing to accept risk-based payment approaches if they were better able to manage patient care. In tiered or broad network products, patients can and will be more likely to seek care from other providers.

As under traditional HMO benefit designs, PCMHs' requiring consumers to share in the costs of services may help reduce the likelihood patients will seek care they do not need. A high deductible health plan (HDHP) could also temper overuse of services from the PCMH, as most services will be subject to the deductible. Primary preventive services, however, would receive first-dollar coverage due to the requirements of the Affordable Care Act (ACA). This could align consumers' and providers' incentives. However, high deductibles might compromise the PCMH's ability to manage the patient's care and to perform well on quality measures, such as those associated with secondary preventive services. Additionally, HDHPs do not work well with primary care capitation. HDHPs generally require a fee-for-service chassis to determine how much of the deductible is subject to the care a patient receives.

Value-based insurance design (V-BID) would support PCMH goals by reducing financial barriers to care for high-value services, thereby aligning patients' interests with the practice's. Additionally, because V-BID lowers consumer cost-sharing for services with well-established, positive effects on the quality of care, the provider can perform well on delivering preventive care—practice patterns that may serve as the basis for bonuses.

Benefit designs that recognize the value of alternative sites of care, such as retail clinics, as covered alternatives to physicians' offices can either support or work against the PCMH, depending on the specific payment approach used. Providers under capitated payments would have a greater incentive to refer to retail clinics for simple preventive and diagnostic services, such as checking for ear infections, whereas providers paid by fee schedule face incentives to perform those services directly, whether convenient for the consumer or not.

If payment migrates from the fee schedule to capitation, whether to support telehealth moves from a benefit design issue to one of practice preference. Under primary care capitation, practices should be open to the robust, lower-cost use of telehealth technologies to communicate with patients and with other health professionals.

Finally, under PCMH payment approaches that place an independent practice at financial risk, the practice would have an interest in assuring that inpatient care recommended by others is really needed. In an HMO environment, the medical home "gatekeeper" typically has control over non-urgent referrals and hospitalizations. That control lessens in a preferred provider organization (PPO), in which patients can gain access to specialists with moderately greater cost-sharing. Further, many hospitalizations and subsequent high-cost procedures occur urgently and outside the PCMH's purview. An effective utilization management program including precertification (i.e., review by the health plan) and continued stay review (if payment is based on per diems rather diagnosis related groups) would support the medical home's objectives regarding both clinical appropriateness and cost control.

Environmental Factors

In general, only HMOs are legally allowed to take on financial risk-sharing, such as primary care capitation, for health care providers. Most states have regulations that ban substantial capitation for providers outside HMOs, because of concern that those providers cannot handle the risk financially and that fixed payments impart incentives for stinting on care. Similarly, Medicare also limits the amount of risk-sharing a Medicare Advantage HMO can take on, having provided detailed guidance on what is considered "substantial financial risk"[1] and therefore not permitted. In short, as of now, substantial use of capitation to support the PCMH would seem prohibited in PPOs.

Although there is hope that innovations riding atop fee schedule-based payments will support medical homes, capitation may better support the medical home concept. Yet, today, primary care capitation is typically used only in HMOs—only HMOs can use a gatekeeper system in which a consumer selects a single primary care practice for services and for access to specialty care.

Conclusion

On the supply side, some payment methods will discourage providers from generating face-to-face visits, shorten their time with individual patients, and encourage the provision of care that was not

previously paid, such as care coordination and communications with patients via e-mail or telephone. On the demand side, benefit designs, such as narrow networks, can encourage consumers to seek care from providers in a medical home. Moderate cost-sharing and value-based insurance design encourage consumers to be more cost sensitive and to seek necessary services. Alternative sites offer consumers cheaper, more convenient places to receive their care than their provider. And utilization management approaches give providers and payers the authority to monitor and manage their patients' care.

Focused Factories—Specialty Service Expertise

The focused factory is characterized by a uniform approach to efficiently delivering a limited set of high-quality services. This typically means that a qualified set of specialists provides care for a procedure or a condition and for related services for which they have great expertise. It is possible to pay these organizations using legacy payments such as fee schedules and diagnosis related groups. Yet, the payment approach best able to support the objectives of a focused factory may be bundled episodes, a coordinated payment to all providers related to a procedure, condition, or treatment across an episode of care. Bundled episodes can be procedure specific, condition specific, or somewhere in between—a group of treatments for a given condition.

We can change supply-side incentives by paying a coordinated risk-based payment to all providers in an episode of care. But we also need to change incentives for consumers to seek care from providers with the greatest expertise, to ensure that they are receiving excellent, cost-effective care. Procedure-based episode payments are narrow and therefore easier to define. Reference pricing can pair well with procedure-based episodes that have well-defined payments. This benefit design can also steer patients to more cost-effective providers. Other benefit designs that create richer coverage for use of centers of excellence give patients incentives to seek care from them. Additionally, value-based insurance design can encourage patients to seek particular services, such as diabetes care, by lowering cost-sharing for the services they need to manage their conditions. All of these components can better support the delivery of specialty services by these providers.

Introduction

The "focused factory" is a concept developed by Wickham Skinner in 1974[2] based on the observation that "simplicity and repetition breeds competence." It refers to provider organizations or their subsets that deliver highly specialized care for a defined and limited group of patients. The idea of providers taking a standardized approach to delivering a limited set of high-quality services efficiently contrasts with delivery-reform approaches like ACOs and PCMHs, which attempt to improve the continuum of care for a general population of patients.

Proponents believe focused factories offer clinical, operational, and financial alignment without the complexity inherent in managing a population's health care needs. The focused factory provides a way for specialists, who so far are not central to ACO and medical home initiatives, to enhance the value of the care they provide while avoiding the data requirements, organizational challenges, and potential for monopolistic behavior that can come with ACOs.

The Basic Payment Approach

Although it is operationally possible to pay a focused factory with legacy payment methods—fee schedules for physicians, per diems or diagnosis related groups (DRGs) for hospitals, and the various approaches used to pay postacute care facilities—these methods provide no incentive for health professionals and provider facilities to decrease fragmentation of patient care, or to mount initiatives that reduce costs and eliminate unneeded services.

The best approach to payment for focused factories may be bundled episode payment. Bundling first links payments that otherwise would be made separately to all of the providers performing services, then extends the period of care covered by payment beyond an individual encounter to the entire episode. More ambitious episodes based around a hospitalization extend beyond the hospital discharge (the end point for a DRG episode) to 30, 60, or more post-hospitalization The Centers for Medicare & Medicaid Services is testing bundled episodes that are mostly triggered by a hospitalization. But they are also testing bundled episode payments for chronic conditions, regardless of whether a hospitalization occurs. In contrast to "procedure-based episodes," which are typically triggered by performance of a procedure, usually hospital based, such as joint replacement surgery, "condition-specific episodes" cover the care delivered by all involved providers for a patient with a particular diagnosis, such as diabetes or ischemic heart disease. A "treatment episode," perhaps for a course of chemotherapy, is an intermediate episode designation. The bundled episode payment method has strengths and weaknesses, with some weaknesses related to implementation challenges.

Each approach poses specific operational issues. Within the prevailing system in which providers generate a claim for services rendered, a major operational challenge is determining which claims are part of a bundle and which should be paid separately; longer episodes generate greater chance of error in the allocation of claims to the bundle. For condition-specific episodes in particular, payers need a clear-cut and reliable approach to determining when an episode is triggered.

Payment for bundled episodes can be retrospective or prospective with financial bonuses or penalties. In a retrospective approach, actual expenditures are reconciled using standard payment amounts from the payer against a target payment amount. If the submitted claims for services are less than the target, a bonus payment is made, with the formula for distribution determined in various ways. If spending exceeds the target, the payer is owed a recoupment. Withholds on the routine payments made to providers participating in the bundle facilitate recoupment. In a prospective payment, a single bundled payment is made to one of the providers, often the hospital in a procedure-specific episode. Physicians and other practitioners and providers submit "no-pay claims" to the payer but are paid by the "convener" provider a previously agreed-upon amount.

Complementary Payment Approaches

Procedure- or treatment-episode-based payments can be combined with condition-specific episode payments, such as when coronary stents are the appropriate treatment for a patient with ischemic heart disease. The procedure-based payment might go to different providers than those taking responsibility for the condition.

Procedure- and condition-specific bundled payments also are compatible with other payment methods that reward prudent health spending, including population health payment approaches. In this case, the ACO-like organization receiving shared risk payments or capitation payments would administer the episode-based payments to constituents of the organization, or possibly to other providers, effectively substituting for the payer. Bundled episodes represent a variation in subcapitation approaches that globally capitated organizations have long used.

Complementary Benefits Designs

For partners receiving a bundled episode payment, a fixed financial target is a strong supply-side approach to constraining spending. Yet, an appealing theoretical advantage of bundled episode payment is that it helps insurers manage financial risk. The narrow nature of a procedure-based payment, such as for a knee replacement, permits consumers to calculate their portion of the bundled price up front and select a provider accordingly. So, benefit designs that, for example, establish a reference price for the procedure-based bundled episode have theoretical appeal. Additionally, reference pricing would channel consumers to lower-cost providers. Condition-based bundled episodes

are complex, with varying services, providers, and episode lengths, which may make reference pricing too difficult to establish.

The practical challenge to this concept, however, is that for virtually all bundled episodes, the patient's annual out-of-pocket maximum protection will kick in during the course of care. Medicare benefits do not include annual out-of-pocket maximums, but more than 90 percent of beneficiaries have supplemental coverage, which does provide such protection. As a result, cost-sharing—whether in the form of reference pricing or high deductibles that are designed partly to make consumers price and cost conscious—likely has greater impact on consumers' choices for discrete, often one-time services, such as colonoscopies or MRI scans, than for care provided over time in an episode. To maintain consumers' incentives for consumers to make high-value choices throughout an episode, continuous, thin cost-sharing—on the order of 5 to 10 percent—might better remind patients that costs are associated with each additional service they receive (de Brantes, Berenson, and Burton 2012).

Centers of excellence (COEs) can complement bundled episodes. In this combination, consumers or Medicare beneficiaries would have financial incentives to seek care from designated centers, which are typically selected because of demonstrated expertise and cost effectiveness in a discrete service line, akin to a focused factory. Quality is always a consideration in the designation of COEs. Purchasers and payers can put COEs under contract to follow evidence-based appropriateness guidelines and can monitor compliance to guard against inappropriate bundles.

Outside of COEs, payers can rely on regular utilization management tools, especially precertification, to protect against paying for unneeded episodes. The problem is that there is often a grey zone around appropriateness. Even external payer-based reviewers may only be able to detect and prevent clear-cut, inappropriate interventions, as they do under legacy payment approaches. The advantage of condition-specific episodes is that their inherent incentives discourage inappropriate procedural interventions. The concern under condition episodes, instead, is that providers would stint on services, especially high-cost procedural interventions. V-BID is meant to lower cost-sharing for high-value services. But so far, its application is mostly to primary and secondary prevention services related to specific conditions, not to procedures, the indications for which depend on clinical detail. For example, cost-sharing would be lowered for a patient with diabetes who wants an eye exam.

A tiered network could incentivize consumers to seek care from focused factories. This design would place the focused factory in the highest tier where consumers have the lowest out-of-pocket costs. Other specialist groups that are not highly specialized would be placed in lower tiers where

consumers have high out-of-pocket costs. A COE for the particular procedure or condition could also be placed in tier 1.

Environmental Factors

Focused factories will work best in markets where there is competition for patients and where acceptable quality can be assured and reasonable prices prevail. The markets also have to be large enough that the provider will have enough experience to achieve excellence and that ancillary services to improve patient outcomes are justified.

Conclusion

For the focused factory, a bundled episode payment approach has potential as the preferred payment method, but operational challenges may limit their application. Current tests of the approach will help determine its future role in payment reform. On the demand side, benefit designs such as differential benefits for selecting COEs and narrow networks encourage consumers to seek care from providers with particular expertise. Reference pricing can both establish well-defined prices for procedure-based bundles and steer consumers to more cost-effective specialty providers. Value-based insurance design can encourage patients to seek particular services, such as primary preventive services or services beneficial for their specific condition, from the providers in the bundle. All these components can better support the delivery of specialty services by providers with the most expertise in a given procedure or condition.

Accountable Care Organizations—Integrated Delivery Systems

Accountable care organizations (ACOs) are groups of physicians and hospitals that share financial and medical responsibility for providing coordinated care, with financial incentives to provide high-quality care and to limit avoidable, unnecessary spending. This concept was promoted by the Affordable Care Act and has since expanded across the public and private sectors. ACOs can be staffed by a hospital or groups of physicians receiving payment for coordinating care across a population of patients. Typically, the ACO's constituent members—physicians and other health professionals, or hospitals for hospital-based ACOs—receive standard payments for services rendered. Payments are made through fee schedules for physicians and per diems or DRGs for hospitals. The ACO itself is typically under either a shared savings or a shared risk arrangement, and savings or overspending are calculated against a target. Some ACOs are paid on per capita basis, such as with capitation.

The Medicare Shared Savings Program (MSSP) ACO, which pays provider groups under a fee-for-service shared savings arrangement, does not incentivize or require that patients seek care from providers within the ACO. This can make it more difficult for providers to take on the financial responsibility associated with the population attributed to them. However, in the commercial market, employers and other payers can use benefit designs to create incentives for consumers to seek care from the ACO. This supports the ACO's incentive to coordinate care and improve patient outcomes. For ACOs that take on financial risk, in the form of a shared risk arrangement or capitation, a narrow network benefit design might best drive consumers to seek care from ACO providers and those providers only. This makes it easier for the ACO to manage and coordinate its patients' care. Value-based insurance design can also encourage consumers to seek clinically beneficial services, thereby improving patient outcomes and helping ACO providers meet quality standards and become eligible for sharing in any savings. Last, alternative, less expensive sites of care can help ACOs deliver care more cost-effectively and reduce the likelihood that consumers will seek care at a level they do not need (e.g., emergency services). All these components can better support the delivery of integrated patient care.

Introduction

ACOs are networks of physicians and hospitals that share financial and medical responsibility for providing patients with coordinated care, along with financial incentives to limit avoidable, unnecessary spending. Typically, ACOs, whether organized and managed primarily by a hospital system or a medical group practice, have a strong base of primary care. The participating providers are collectively accountable for both the quality and the full cost of care for a population of patients who, in various ways, are assigned to the ACO to make providers accountable. A core objective is to reduce the trend in costs; however, the ACO is typically required to achieve targets of performance on quality measures to be eligible to share savings from any reduction in costs.

The ACO concept is recent and was first promoted by specifications in the ACA. However, commercial insurers have long supported ACO-like organizations with payment and benefit design approaches that support the goals of improved quality and, especially, reduced costs. Organizational structures consistent with the ACO delivery concept include multispecialty group practices, integrated delivery networks, and independent practice associations. In contrast to Medicare, which does not modify its standard benefit structure with prescribed beneficiary cost-sharing obligations, private payers typically develop benefit design approaches that provide consumers with financial incentives to seek care from ACO providers.

Basic Payment Approaches

The ACA prescribed Medicare's approach to ACO development with a specific payment approach along with the standard Medicare benefits package. In general, the various approaches to paying ACOs are referred to as population health methods, because the base or incremental payment is based on the number and characteristics of the individuals assigned to the ACO, without regard to the specific services they receive. Shared savings, the most common payment method established for ACOs, maintains legacy approaches to the constituent providers based on service activity and provides incremental bonuses based on spending for the assigned population as compared against a target spending amount. Shared risk builds on this approach, creating penalties for overspending as well and usually awarding greater amounts for savings. Alternatively, various forms of capitation alter the base payment method to per capita spending rather than use the legacy volume-based payment approaches.

MSSP relies on a payment method called shared savings, under which ACOs receive financial bonuses if their assigned beneficiaries' health care costs are below a projected target amount, which is

based on the ACO providers' historic spending. (This approach is sometimes referred to as "upside-only" shared savings.) In this approach, the provider members of the ACO receive their usual fee schedule or diagnosis-related group payments; the ACO entity itself is eligible for shared savings if it reduces spending to a minimum savings threshold (to account for normal variations in health care spending), contingent on its performance on quality measures. MSSP ACOs can choose to participate in two-sided, shared risk arrangements—in addition to receiving bonuses, ACOs with expenditures at or above a minimum loss threshold have to repay excessive spending. Similarly, Pioneer ACOs, which are Medicare demonstrations, can be paid using a variety of population-based methods, including partial or global capitation (i.e., per capita payments per month).

Commercial ACOs are often paid shared savings based on historic costs trended forward, although there is variation. An increasing number of commercial ACOs are being established with shared risk arrangements or are entering into contracts that migrate the payment method to shared risk over three years.

Complementary Payment Approaches

One challenge for shared savings and shared risk ACOs is that the constituent providers continue to face volume-enhancing payment incentives through standard payment methods, such as physician fee schedules and DRGs for inpatient care. Indeed, two-sided risk is similar to global capitation in how it penalizes the ACO for unnecessary spending. Yet, the legacy payment methods that represent constituent providers' cash flow may overwhelm ACOs' incremental shared savings incentives for more prudent spending. It is for that reason that some consider shared risk an "on-ramp" to prepayment through capitation.

Capitation provides front-end capital to the ACO, permitting the entity to employ payment methods for its constituent providers that deviate from standard payment, with greater incentives for reduced spending. That is, the capitated ACO can use primary care and specialty capitation to distribute the revenues it receives from payers to providers in the ACO. Similarly, the capitated ACO may be in a better position than a more removed third party to pay based on procedure- or condition-based bundles, as it can better prevent potential unintended behavioral responses. For example, given the concern that procedure-based bundled episode payment remains volume based (i.e., providers still have incentives to do more bundled episodes), the ACO can assure that providers follow evidence-based guidelines for appropriate care. Similarly, the ACO is in a better position than the insurer to guarantee

that condition-specific episodes of care meet the clinical criteria required to be eligible for a bundled payment.

Commercial insurers have increasingly adopted DRGs as the payment method of choice for inpatient care, to be consistent with Medicare and to transfer risk for an entire hospital stay, rather than each day, to the hospital (except for outlier cases). However, under population-based payment methods that place ACOs at risk for the individuals assigned to them, the ACO might actually prefer to maintain the risk itself rather than pass it on to the hospital: it can actively manage whether the patient gets admitted, and it can better assure a high-quality, "early" discharge that includes follow-up when the patient returns to the community or to a postacute care facility. In short, using per diems can benefit an at-risk ACO with regard to early discharge, rather than using DRGs to give the hospital the savings.

This example emphasizes that hospital-based ACOs and physician-based ACOs may have different perspectives on which payment methods to employ, if they can receive capitated payments and then pay participating providers. In particular, assuming a reasonably competitive market for hospital services, a physician-based ACO may be in a position to exert negotiating leverage with competing hospitals, achieving more favorable prices and better terms and conditions than a hospital-based ACO committed to its own hospital and other facilities.

Complementary Benefits Designs

ACOs, whether receiving base capitation payments or incremental shared savings or shared risk, would be supported by narrow and tiered network benefit designs. The objective would be for favorable cost-sharing to encourage consumers to use ACO providers. Presumably, the payer has achieved some price concessions from the ACO constituent providers in exchange for increased volume. Further, if selection into the favored tier or narrow network is determined by performance-based quality and service-use benchmarks, the ACO would improve its value of care by channeling patients to its providers through differential patient-cost-sharing.

From the ACOs perspective, a narrow network design would be preferable to a tiered network: ACO providers would have more control over clinical care decisions, including referral preferences, because patients have more limited choice of providers in the narrow network. Providers in a tiered network would likely be less willing to assume risk, because their patients still have access to providers in all tiers. However, a shared savings (upside-only) arrangement, could work well with a tiered network by aligning patients' outcomes to providers' performance and eligible savings.

HDHPs may create a mixed set of interactions with population-based payment methods in which providers bear financial risk. On one hand, because providers are bearing risk, their patients face a high deductible; this has the direct effect of influencing patients to seek less care, which reduces spending to the ACO's benefit. On the other hand, ACOs, typically based on primary care practices, seek more management control over patient care; therefore, financial barriers to care (through a high deductible) can interfere with patients' acceptance of ACO providers' judgment. ACOs need to meet quality standards and, for that reason alone, would want to see lowered barriers to care. Otherwise, adverse outcomes will occur when patients choose to disregard their physicians' advice. HDHPs typically provide first-dollar coverage for primary preventive services, which eases the tension between patients' incentives and providers' directions. However, secondary preventive services and other clinical services that might affect more serious conditions and complications are subject to the full impact of the high deductible, for good and bad.

Similar to the impact of waving cost-sharing for primary prevention inside a HDHP, value-based insurance design complements to population-based payment in how it can encourage consumers to seek preventive services. V-BID reduces financial barriers to seeking care and can identify services of high value, such as maintenance medications for chronic conditions. This adds an additional demand-side tool to the delivery-side population-based payment approach. V-BID could lower consumer cost-sharing for secondary preventive or other clinical services subject to a deductible. But for an HDHP with a health savings account, according to IRS standards, all services must be subject to the deductible regardless of other incentive designs. V-BID could encourage nuanced fee-for-service payment, so that providers performing clinically beneficial procedures receive regular payments and those providing unnecessary services do not.

Both HDHPs and V-BID were designed to complement standard methods for paying health professionals and hospitals. As a result, they are both readily compatible with a shared savings or shared risk approach to population-based payment, as these approaches rely directly on usual fee schedule and per diem or DRG hospital payments. The retrospective determination of shared savings or losses does not affect patients' cost-sharing obligations or any impact cost-sharing might have on their care-seeking behavior. HDHPs are not readily compatible with professional or global capitation because they do not include the usual fee-for-service claims on which a patient's cost-sharing is calculated. Traditionally, co-payments, rather than co-insurance or deductibles, have been applied when payment to the integrated provider entity is made by capitation. Therefore, V-BID can be compatible with capitation if co-payments are the cost-sharing mechanism used to encourage certain

care-seeking behavior. In addition, HDHPs could function with capitation if the plan administers deductibles based on each patient encounter, though this is operationally difficult to do.

Finally, benefit designs that encourage consumers to seek care from alternative sites have potential strengths and weaknesses in a non-risk-bearing, fee-for-service environment. Telehealth might be susceptible to overuse if the providers employing it are not sharing any risk. However, risk-bearing ACOs would have strong interest in preventing frivolous or otherwise inappropriate use of alternative sites of care, including telehealth. Similarly, a risk-bearing ACO might wish to contract with, manage, or own retail or workplace clinics as well as telehealth services, as a way to facilitate care in environments with lower cost than emergency rooms. In this way, promoting alternative sites of care and population-based payments are complementary.

Environmental Factors

A fundamental factor in commercial insurers' expanding the use of global payment approaches to support ACOs lies in how states regulate risk-bearing entities. Today, most state regulations only allow the various forms of capitation, ranging from primary care capitation to full global capitation, to take place in licensed HMOs and not PPOs or other unlicensed products. The theory is that HMO requirements and oversight provide both needed protection against stinting on care and the ability to take on insurance risk. With current trends continuing to move away from HMOs toward PPOs, fostered by the growth of self-funded, employer sponsored insurance not subject to state regulation, prepaid capitation may have less potential as a payment model for ACOs.

Less clear is the degree to which shared risk will be permitted on a PPO commercial insurance platform. While Medicare payment methods are not subject to state regulation, states might regulate how commercial payers set up both shared risk payment arrangements and provider networks. States may only tolerate shared risk up to a certain limit or allow providers to apply losses to future management fees rather than to out-of-pocket payments to the insurance entity. States may apply regulations similar to those for capitation arrangements that require significant insurance risk by an unlicensed entity.

Conclusion

With a successful combination of incentives for consumers and health care providers, integrated care delivery could deliver better health care and create healthier populations. On the supply side, potential payment methods could encourage providers to coordinate care by offering incremental bonuses or penalties for care that meets both cost and quality targets, perhaps as a transitional payment method leading to population-based payment methods like capitation. On the demand side, benefit designs, such as narrow networks, can encourage consumers to seek care from providers in an ACO. Value-based insurance design can encourage consumers to follow clinical guidelines and seek evidence-based care that can lead to positive outcomes and therefore, savings for providers. And alternative sites of care can help providers operating under shared risk payment arrangements provide cheaper, more convenient care to consumers who do not need care in traditional settings. All of these components can better support the delivery of integrated patient care.

Notes

1. Federal Register Vol. 63, No. 123, Rules and Regulations. 42 CFR 422.208/210, June 26, 1998
2. Skinner, Wickham. "The Focused Factory." *Harvard Business Review*, May 1974, https://hbr.org/1974/05/the-focused-factory.

References

Chen, Melinda A., James P. Hollenberg, Walid Michelen, Janey C. Peterson, and Lawrence P. Casalino. 2011. "Patient Care outside of Office Visits: A Primary Care Physician Time Study." *Journal of General Internal Medicine* 26 (1): 58–63.

de Brantes, Francois, Robert Berenson, and Rachel Burton. 2012. "Payment Reform: Bundled Episodes vs. Global Payments: A Debate between Francois de Brantes and Robert Berenson." Washington, DC: Urban Institute.

Gilchrist, Valerie, Gary McCord, Susan Labuda Schrop, Bridget D. King, Kenelm F. McCormick, Allison M. Oprandi, Brian A. Selius, et al. 2005. "Physician Activities during Time Out of the Examination Room." *Annals of Family Medicine* 3 (6): 494–499.

About the Authors

Robert A. Berenson joined Urban as an Institute fellow in 2003. In this position he conducts research and provides policy analysis primarily on health care delivery issues, particularly related to Medicare payment policy, pricing power in commercial insurance markets, and new forms of health delivery based on reinvigorated primary care practices. In 2015, Berenson was appointed to the Physician-Focused Payment Model Technical Advisory Committee set up by Congress through the Medicare and CHIP Reauthorization Act of 2015 to advise the Secretary of the Department of Health and Human Resources. In 2012, Berenson completed a three-year term on the Medicare Payment Advisory Commission, the last two years as vice chair. From 1998 to 2000, he was in charge of Medicare payment policy and private health plan contracting in the Centers for Medicare and Medicaid Services. Previously, he served as an assistant director of the White House Domestic Policy Staff under President Carter. Berenson is a board-certified internist who practiced for 20 years, the last 12 years in a Washington, DC, group practice. He is a graduate of the Mount Sinai School of Medicine and a fellow of the American College of Physicians.

Suzanne F. Delbanco is the executive director of Catalyst for Payment Reform (CPR), an independent, nonprofit corporation working on behalf of large health care purchasers to catalyze improvements to how we pay for health services and to promote better and higher value care in the United States. In addition to her duties at CPR, Delbanco serves the board of the Health Care Incentives Improvement Institute. Previously, Delbanco was the founding chief executive officer of The Leapfrog Group. Delbanco holds a PhD in public policy from the Goldman School of Public Policy and an MPH from the University of California, Berkeley School of Public Health.

Roslyn Murray is the project and research assistant for CPR. In this role, Murray provides background support and conducts research for CPR on policy and health care payment reform issues. She has also held research assistant positions at Stanford School of Medicine, the National Institutes of Health's National Heart, Lung and Blood Institute, and San Francisco General Hospital. A graduate of Stanford University, Roslyn holds a bachelor's degree in human biology, health, and health policy.

Divvy K. Upadhyay is a research associate in the Health Policy Center at the Urban Institute, where he focuses on qualitative research on health reform measures at the federal and state levels. His policy research interests include reforms related to physician and hospital payment; health care delivery; the Affordable Care Act; Medicare and Medicaid; and issues in medicine related to primary care and

diagnosis errors. Before joining Urban, Upadhyay worked with the editorial team of *Health Affairs*. Trained as a physician, Upadhyay holds an MPH with a focus on health policy and organization from the University of Alabama at Birmingham.

Statement of Independence

The Urban Institute strives to meet the highest standards of integrity and quality in its research and analyses and in the evidence-based policy recommendations offered by its researchers and experts. We believe that operating consistent with the values of independence, rigor, and transparency is essential to maintaining those standards. As an organization, the Urban Institute does not take positions on issues, but it does empower and support its experts in sharing their own evidence-based views and policy recommendations that have been shaped by scholarship. Funders do not determine our research findings or the insights and recommendations of our experts. Urban scholars and experts are expected to be objective and follow the evidence wherever it may lead.

2100 M Street NW
Washington, DC 20037

www.urban.org