THE HAMILTON PROJECT

Proposal 1: Transitioning to Bundled Payments in Medicare

Michael Chernew
Harvard Medical School

Dana Goldman
University of Southern California

Deficit Reduction (10-year): $100 billion

Broader Benefits: Promotes efficiency in the Medicare program by providing incentives to treat disease rather than paying for individual services; continues to encourage improvements in the quality of care, but at lower costs.

Introduction

The core challenge for the health-care system in general—and for care financed by the federal government through Medicare—is how to reduce inefficient spending while continuing to improve the quality of care. The fee for service (FFS) system, as currently practiced, is at the heart of this challenge, particularly for Medicare, because it is focused on providing and paying for medical services rather than on promoting and incentivizing medical outcomes.

The existing FFS portion of Medicare, which enrolls almost 75 percent of Medicare beneficiaries, relies on a byzantine system of fee schedules. There are thousands of codes for different services; setting the appropriate fee is enormously complex. Mispriced fees create incentives leading to the overuse (or underuse) of medical services. As a result, resources flow to overpriced activities and infrastructure. Importantly, the FFS system reduces incentives for providers to be efficient over the entire episode of care (Chernew, Frank, and Parente 2012; Landon 2012).

We propose a strategy for transitioning away from FFS payment to a global payment model. These changes are designed to promote efficiency in the Medicare program and facilitate the ability of health-care providers to continue to improve the quality of care, but along a dramatically slower spending trajectory. Compared to likely budget scenarios outlined in the Congressional Budget Office's (CBO's) "Alternative Fiscal Scenario," for example, this proposal would reduce Medicare outlays within the ten-year budget window by more than $100 billion; by reducing the growth in health spending, it would reduce spending in later years (CBO 2013). Moreover, this proposal could contribute to lower outlays for other government-financed health care, and to improvements in the provision of health care in the economy at large. In particular, we support three proposals:

1. The Medicare program should create a global payment model (that can operate independently from the existing FFS system), in which provider systems are paid a fixed fee (or given a fixed budget) to cover all beneficiary spending.

2. Congress and Centers for Medicare and Medicaid Services (CMS) should create regulatory neutrality between Medicare Advantage (MA) plans and accountable care organizations (ACOs).

3. Congress and the CMS should create a safe haven from regulations if an organization accepts global payment.

DISCLAIMER: The views expressed in this paper are solely those of the authors and not of any institutions or government agencies with which they are or have been affiliated.

There are many advantages to transitioning toward a global payment model. One important advantage is that such a model orients the incentives of providers toward taking advantage of efficiencies. It should be noted that a global payment model will likely require organizational changes; there are some concerns that it may lead to lower-quality care. We believe, however, that these concerns can be addressed within the system we advocate.

HEALTH-CARE TERMS[1]

Accountable Care Organization (ACO): An ACO is a group of health-care providers who give coordinated care and chronic disease management, and thereby improve the quality of care for patients. The ACO's payment is tied to achieving health-care quality goals and outcomes that result in cost savings.

Fee for service (FFS): FFS is a method in which doctors and other health-care providers are paid for each service performed. Examples of services include tests and office visits.

Global Payments System: As opposed to FFS, this is a method in which doctors and other health-care providers are paid a fixed fee to cover all beneficiary spending.

Medicare Part A: Hospital insurance paid for by a portion of the Social Security tax. It helps pay for inpatient hospital care, skilled nursing care, hospice care, and other services.

Medicare Part B: Medical insurance paid for by the monthly premiums of people enrolled, as well as by general government funds. It helps pay for doctors' fees, outpatient hospital visits, and other medical services and supplies that are not covered by Part A.

Medicare Part C (Medicare Advantage, or MA): A type of Medicare health plan offered by a private company that contracts with Medicare to provide Part A and Part B benefits.

Medicare Part D: Prescription drug coverage that is voluntary and paid for by the monthly premiums of enrollees and Medicare.

The Challenge

The American health-care system is enormously complex. To guide what can often be an esoteric discussion, the box below contains some important terms and definitions related to the health-care sector.

Medicare's unmanaged FFS system is particularly convoluted, to say the least. While many organizations use FFS within settings with other tools to affect behavior, Medicare's approach has led to considerable inefficiencies (Reinhardt 2012). There is a separate fee schedule for every type of nondrug provider. All these schedules are complicated. The physician payment system has more than 7,000 codes for unique services. Many are variants on the same type of service. For example, there are ten codes for physician office visits that vary based on new versus established patients, and on the level of complexity. There are rules to define each visit; using these rules, providers assign a complexity level based on time spent or the nature of the visit. For example, a level-three office visit is one that requires at least two of the following three components: an expanded problem-focused history, an expanded problem-focused examination, and medical decision making of low complexity. For other services, there are also multiple codes: there are more than forty codes for CT scans based on the part of body scanned and which type of contrast agent is used. This level of intricacy pervades the system.

Moreover, setting the appropriate fees is thorny because the fee schedule must adjust for economies of scope (scanning two body parts in one sitting should cost less than twice scanning a single body part). Assumptions about capacity utilization, which may vary in different settings, and the lifetimes of high-cost equipment are needed to set an appropriate fee.

The existing set of fees is clearly flawed. The variation in fees for any given service based on the setting of care is almost surely wider than can be justified, although some variation may be appropriate. The process for setting physician fees (a process that relies heavily on recommendations from committees of physicians) and facility fees is cumbersome and widely criticized for favoring specialties over primary care. For example, under this system primary physicians are paid considerably less per hour for cognitive services than specialists are paid for procedures (Bodenheimer, Berenson, and Rudolf 2007).

The problems that arise because of mispriced services extend beyond simple inequitable allocation of funds across providers. Mispriced fees create incentives that result in overuse (or underuse) of medical services. They incent resources to flow to overpriced activities and infrastructure and away from underpriced activities and infrastructure.

Perhaps the most important point is that the FFS system diminishes incentives for providers to be efficient over the entire episode of care (i.e., for all of the care associated with the treatment for a particular problem or condition over a period of time). If hospitals work to reduce readmissions, they lose income. If physicians reduce unnecessary office visits, they lose income. While undoubtedly providers strive to provide high-quality care, in a purely economic sense it is difficult for them to justify devoting resources to reducing use of unnecessary services or to finding less-resource-intensive ways to deliver an episode of care.

We therefore believe that the Medicare payment system is ripe for reform. The proposal might not save considerable budget savings when judged against current law (the bar against which the fiscal consequences of payment reform is judged) because that trajectory is at a historic low. In that context, it could be interpreted as improving health outcomes within the current cost structure. However, this current law baseline includes cuts to physician payments that policymakers have been reluctant to implement in the past, cuts that, if not implemented, would increase deficits by more than $100 billion over ten years. Hence, an alternative interpretation is that the proposal could contribute about $100 billion of deficit reduction that policymakers would otherwise have needed to find elsewhere.

The Proposal

Our proposal is based on three provisions:

1. **The Medicare program should create a global payment model (that can operate independently from the existing FFS system), in which provider systems are paid a fixed fee (or given a fixed budget) to cover all beneficiary spending.**

   The global payment is similar to both the global budget used in the existing pioneer ACO model and the per beneficiary premium contribution used in the MA plan (Chernew, Frank, and Parente 2012). Specifically, under our model, the CMS will pay a health plan or provider system a fixed payment (or set a fixed budget) to cover all medical services, including prescription drugs. As described below, MA plans and ACOs will be able to charge beneficiaries a premium above the global rate that represents the federal contribution.

   Unlike the existing ACO and MA models, under our program, the fixed payment will be set in a manner independent of the existing FFS system. Under the existing system, the operation of the ACOs and MA plans rely on the existence of the FFS program, which, if ACOs and MA plans are successful, will wither.

   The budget effects of a global payment model depend on the global payment rates. Setting the global payment is a political decision. We advocate, as a default, that the payment be set to match the current law, per beneficiary Medicare spending, and that it rise at the rate of the current law per beneficiary spending trajectory. This trajectory is rising even after inflation adjustment, but at a much slower rate than rates in the past. Thus, a revised fee trajectory that will allow inflation adjusted per beneficiary payment to rise at the same rate as current law (about 1.1 percent per year) could be developed so that the ten-year budget score remains the same. This is equivalent to about 0.7 percentage points less than GDP growth over the next ten years. Congress could always modify the global rate as it does with the existing fee schedule, but we propose any changes be implemented with a three-year lag to provider plans or providers, with certainty about the target and assurances that efficiency gains will not be captured by the government via lower rates the following year.

2. **Congress and the CMS should create regulatory neutrality between MA plans and ACOs. In a global payment model, the payment can go either to a health plan, as in the MA program, or to a provider system, as in the ACO program. Regulation should strive to level the playing field between these two organizational forms.**

   Most importantly, this means that the payment rates for MA plans and ACOs should be equivalent. Accomplishing this equivalence will require attention because MA payment rates are set based on county spending, and ACO rates are based on delivery system specific spending. A transition period will be needed, but we believe that ultimately we should move to payment rates that are adjusted for case mix and differences in input costs across areas, but not rates that are reflective of different practice styles across delivery systems or geographies.

   Other areas of regulation should be examined as well. For example, MA plans currently control benefit design and can use that authority to implement value-based insurance design plans, which align copays with the value of medical services. They also can use benefit design to incent beneficiaries to use preferred providers. ACOs do not have this authority at this time. Allowing ACOs to have such authority would address concerns about leakage, but might require other changes, such as having beneficiaries proactively select their ACO as opposed to being assigned by the CMS to an ACO without their knowledge.

   Moreover, in MA, plans bid relative to an administratively set benchmark. If they bid below the benchmark, plans can offer more-generous benefits or rebate Part D or Part B premiums. They can offer even-more-generous benefits if they charge an additional premium. If they bid above the benchmark, they must charge a premium for the standard benefit. They can offer additional benefits if they charge an additional premium. ACOs do not have that freedom. Allowing them such flexibility would allow ACOs that are particularly efficient to attract more beneficiaries, and allow those that are higher quality to charge for any added expense. Perhaps both of those objectives can be met if ACOs establish their own MA plans, but there

are regulatory hurdles to that strategy. Other differences between the programs, such as degree of risk sharing, also exist. Total uniformity is not needed and some heterogeneity may be desirable, but regulation should not favor one organizational form over another.

In both programs, policymakers must be concerned with market power, with fees charged by MA plans or ACOs above the global rate, and with fees providers charge to health plans. For example, caps of additional premiums that could be charged might be imposed to address MA premiums and ACO supplemental fees. These caps could be related to measured quality.

Addressing antitrust concerns in the market for health-care services (e.g., what providers charge MA plans) is more complex because of the vast number of services being purchased and the variation in how providers and plans contract (e.g., diagnosis-related group [DRG] versus per diem). Because integration of care may generate efficiencies, regulatory response to market power might focus on price regulation as opposed to breaking up delivery organizations. Limits on the ratios between negotiated fees charged to MA plans and Medicare rates may be needed, but as the FFS system withers, this approach will not be sustainable and other benchmarks (such as national average prices) will be needed.

3. **Congress and the CMS should create a safe haven from regulations if an organization accepts global payment.**

Many regulations in Medicare are designed to prevent overutilization of care incented by the FFS system. These include regulations against self-referral, various caps on service use, or required utilization review for services such as occupational therapy. In a global payment model, these incentives are eliminated. As a result, they simply represent administrative inefficiencies and needless restrictions. Thus, organizations accepting global payment may be exempt from such rules.

ADVANTAGES TO THIS PROPOSAL

The fundamental challenge facing Medicare is how to slow the rate of growth in public spending while still providing needed access to care for beneficiaries and sufficient resources for providers. The FFS system is an impediment to achieving that goal. The spending trajectory that exists under current law, dominated by FFS, sets ambitious goals; many have questioned whether those goals can be sustained. More important, the FFS system does not allow providers to capture savings from efficiencies they may achieve. This reduces incentives to invest in finding such efficiencies. A global payment model provides such incentives. Similarly, a global payment model also encourages providers to direct care to the most efficient setting as opposed to exploiting differential payment across settings in the current system. Moreover, a global payment model can eliminate the need for some intrusive regulations. Finally, a global budget model provides predictability in spending and spending growth.

Yet despite these advantages, we recognize a number of challenges exist. The most important point here is that success under a global payment model likely requires organizational change. Many providers may not be ready to accept the risk inherent in global payment. By keeping the current system as a fall back, providers will not be forced into the global payment model. Of course, these organizations may not fare well in the existing system with the current schedule of fee updates. As payment rates fail to keep up with input price inflation, they will face financial distress, so relative to current law global payment may be appealing. In Massachusetts, diffusion of global payment was very rapid, and included practices that were not part of large integrated systems. Furthermore, under the global payment model we propose, inflation-adjusted payment rates rise each year, suggesting organizations do not need to reduce spending to be successful: they only must control the rate at which spending increases.

We also recognize that even in a global payment model there will be uses for FFS and FFS-type systems. For example, risk adjustment may require continued collection of service-level data that may use the existing coding system, though perhaps modification can be made as the purpose changes (Ginsburg 2012). Moreover, within provider organizations (or between MA plans and providers) a payment system will be needed (Landon 2012). That system may have aspects of FFS (e.g., bonuses for productivity), but the decisions about how to balance financial incentives with other managerial techniques will reside with the organization, not the government.

Another concern is that such models will encourage health-care systems to provide poor-quality care. Evidence from the 1990s is that while HMOs do not uniformly lead to worse quality of care, elderly and chronically ill patients enrolled in HMOs had worse quality-of-care outcomes than their FFS counterparts (Miller and Luft 1997). Existing evidence from newer models suggests that such models may improve some aspects of quality (Song et al. 2011). Yet quality measures are imperfect and these concerns about adverse effects on quality are genuine. Greater development of quality measurement systems is required.

There are several reasons to believe quality concerns can be mitigated. For example, because payment rates would be rising, the financial resources exist to provide ever-improving quality, and efforts to eliminate waste may actually improve

quality in some cases (because incentives would encourage providers to reduce the rates of complications and delivery of needless services, some of which could have adverse side effects). Furthermore, by allowing providers to charge above the global rate, those consumers who wish to pay more for better-perceived quality will be allowed to do so.

The potential for copremiums or surcharges above the global rate raises another concern about the impact on disparities in access. Lower-income individuals will be less able to buy access to potentially higher-quality systems. Quality measurement systems can be used to create a minimum standard, but again, such systems are inevitably imperfect. Therefore disparities must be monitored and policymakers may need to develop systems to protect low-income beneficiaries. But it is useful to note that under the current system, lower cost is not synonymous with lower quality; it may be the case that a global payment system, with plans or delivery systems accountable for outcomes, provides even better quality for low-income beneficiaries. Thus, relative to the status quo, this proposal may be an improvement.

A final concern is that the system we propose does not save any money relative to the status quo. Under current law, inflation-adjusted spending per beneficiary is forecast to rise at historically low rates (0.7 percentage point below GDP growth compared to an average of 1.5 percentage points above GDP since 1985) (CBO 2012). Under our proposal, policymakers could opt for lower spending targets, but we consider the existing current law trajectory to be sufficiently ambitious. It should be noted that relative to the alternative fiscal scenario, this plan would reduce spending by about $100 billion over ten years.

It is important to recognize that, under a global payment model, further savings can only be captured by the government if the global payment is reduced. For example, reductions in benefit generosity only save money for Medicare if the global payment rates are adjusted accordingly. Regardless of whether Medicare sets the payment to reduce spending relative to current law, our proposal focuses on transforming the incentives in Medicare to encourage efficiency and render the existing spending forecasts more feasible.

Conclusion

The Medicare program is in dire need of payment reform. The FFS system is difficult to manage, lacks incentives for the delivery system to invest in achieving efficiencies, and has historically encouraged unsustainable spending growth. We propose replacing the FFS system with a global payment model designed to limit public expenditures to the current law trajectory, which would represent a significant departure from past levels of spending growth.

Many of the structures needed to implement our proposals exist under current law, including the MA and ACO programs. Yet our proposal differs from current law in three important ways: First, we break the tie between payment and FFS spending. Second, we strive to harmonize the ACO and MA programs. Third, we create regulatory safe havens for organizations accepting global payment. Certainly the challenges to such a migration are great, but the alternative—payment rates that statutorily rise at rates below the rate of inflation in input prices and offer no incentives for efficiency—seems even less appealing.

Authors

Michael Chernew

Professor of Health Care Policy, Department of Health Care Policy, Harvard Medical School

Michael Chernew, Ph.D., is a professor in the Department of Health Care Policy at Harvard Medical School. Chernew's research focuses most notably on the causes and consequences of growth in health-care expenditures, geographic variation in medical spending and use, and value-based insurance design. Chernew is a member of the Medicare Payment Advisory Commission, the Congressional Budget Office's Panel of Health Advisers, and Commonwealth Foundation's Commission on a High Performance Health System. In 2000, 2004, and 2011, he served on technical advisory panels for the Center for Medicare and Medicaid Services. He coedits the *American Journal of Managed Care* and is a Senior Associate Editor of *Health Services Research*. In 2010, Chernew was elected to the Institute of Medicine (IOM) of the National Academy of Sciences, and serves on the Committee on the Determination of Essential Health Benefits. Chernew earned his undergraduate degree from the University of Pennsylvania and a doctorate in Economics from Stanford University.

Dana Goldman

Professor and Leonard D. Schaeffer Chair, University of Southern California

Dana Goldman is a Professor and the Leonard D. Schaeffer Chair at the University of Southern California, where he is the founding director of the Leonard D. Schaeffer Center for Health Policy and Economics. Prior to this appointment, he held the Chair in Health Economics at the RAND Corporation; was the first director of RAND's program in Health Economics, Finance, and Organization; and was the director of the Bing Center for Health Economics. Goldman's research focuses on health economics, Medicare, aging, disability, and chronic disease. Goldman's recognitions include membership in the Institute of Medicine, the inaugural MetLife Foundation Silver Scholar Award, the Eugene Garfield Economic Impact Prize, and the National Institute for Health Care Management's Alice S. Hersh New Investigator Award. He is a health policy adviser to the Congressional Budget Office, and has been featured in numerous media and leading medical, economic, health policy, and statistics journals. He received his B.A. summa cum laude from Cornell University and a Ph.D. in Economics from Stanford University.

Endnotes

1. Definitions are based on U.S. Department of Health and Human Services (n.d.) and U.S. Social Security Administration (n.d.).

References

Bodenheimer, T., R. A. Berenson, and P. Rudolf. 2007. "The Primary Care-Specialty Income Gap: Why it Matters." *Annals of Internal Medicine* 146 (4): 301–306.

Chernew, M. E., R. G. Frank, and S. T. Parente. 2012. "Slowing Medicare Spending Growth: Reaching for Common Ground." *American Journal of Managed Care* 18 (8):1–6.

Congressional Budget Office (CBO). 2012. "The 2012 Long-Term Budget Outlook." http://www.cbo.gov/publication/43288

———. 2013. "The Budget and Economic Outlook: Fiscal Years 2013 to 2023." http://cbo.gov/sites/default/files/cbofiles/attachments/43907_Outlook_2012-2-5_Corrected.pdf

Ginsburg, Paul B. 2012. "Fee-For-Service Will Remain a Feature Of Major Payment Reforms, Requiring More Changes In Medicare Physician Payment." *Health Affairs* 31 (9): 1977-1983.

Landon, Bruce E. 2012. "Keeping Score under a Global Payment System." *New England Journal of Medicine* 366 (5): 393–395.

Miller, R. H., and H. S. Luft. 1997. "Does Managed Care Lead to Better or Worse Quality of Care?" *Health Affairs* 16 (5): 7–25.

Reinhardt, U. E. 2012. "How Medicare Is Misrepresented." *Economix: New York Times*. December 7. http://economix.blogs.nytimes.com/2012/12/07/how-medicare-is-misrepresented/

Song, Z., D. G. Safran, B. E. Landon, Y. He, R. P. Ellis, R. E. Mechanic, M. P. Day, and M. E. Chernew. 2011. "Health Care Spending and Quality in Year 1 of the Alternative Quality Contract." *New England Journal of Medicine* 365 (10): 909–918.

U.S. Department of Health and Human Services (DHHS). n.d. "Glossary." http://www.healthcare.gov/glossary/

U.S. Social Security Administration. n.d. "Differences between Medicare Parts A, B, C and D." http://ssa-custhelp.ssa.gov/app/answers/detail/a_id/167/~/differences-between-medicare-parts-a,-b,-c-and-d

15 Ways to Rethink the Federal Budget