Global Budgets and Technology-Intensive Medical Services

Zirui Song, Ph.D.[1,2], **A. Mark Fendrick, M.D.**[3], **Dana Gelb Safran, Sc.D.**[4,5], **Bruce Landon, M.D., M.B.A.**[1,6], and **Michael E. Chernew, Ph.D.**[1,2]

[1]Department of Health Care Policy, Harvard Medical School

[2]National Bureau of Economic Research

[3]Departments of Internal Medicine and Health Management & Policy, Center for Value-Based Insurance Design, University of Michigan

[4]Blue Cross Blue Shield of Massachusetts

[5]Department of Medicine, Tufts University School of Medicine

[6]Division of General Medicine and Primary Care, Department of Medicine, Beth Israel Deaconess Medical Center

Abstract

Background—In 2009-2010, Blue Cross Blue Shield of Massachusetts entered into global payment contracts (the Alternative Quality contract, AQC) with 11 provider organizations. We evaluated the impact of the AQC on spending and utilization of several categories of medical technologies, including one considered high value (colonoscopies) and three that include services that may be overused in some situations (cardiovascular, imaging, and orthopedic services).

Methods—Approximately 420,000 unique enrollees in 2009 and 180,000 in 2010 were linked to primary care physicians whose organizations joined the AQC. Using three years of pre-intervention data and a large control group, we analyzed changes in utilization and spending associated with the AQC with a propensity-weighted difference-in-differences approach adjusting for enrollee demographics, health status, secular trends, and cost-sharing.

Results—In the 2009 AQC cohort, total volume of colonoscopies increased 5.2 percent (p=0.04) in the first two years of the contract relative to control. The contract was associated with varied changes in volume for cardiovascular and imaging services, but total spending on cardiovascular services in the first two years decreased by 7.4% (p=0.02) while total spending on imaging services decreased by 6.1% (p<0.001) relative to control. In addition to lower utilization of higher-priced services, these decreases were also attributable to shifting care to lower-priced providers. No effect was found in orthopedics.

Conclusions—As one example of a large-scale global payment initiative, the AQC was associated with higher use of colonoscopies. Among several categories of services whose value may be controversial, the contract generally shifted volume to lower-priced facilities or services.

Address correspondence to Zirui Song, Department of Health Care Policy, Harvard Medical School, 180 Longwood Avenue, Boston, MA 02115. zirui_song@hms.harvard.edu..

INTRODUCTION

The growth of health care spending is a major policy concern.[1234] Over the past few years, insurers have begun to adopt new payment strategies centered on moving away from fee-for-service towards bundled payments.[5] Global budgets, the most inclusive form of bundled payment in which groups of physicians and hospitals—increasingly in the form of accountable care organizations (ACOs)—receive a fixed amount for all of a patient's care over a defined time period, are currently being implemented by Medicare and private insurers across the country.[67] ACOs share savings if total medical spending for their patient population comes in under the budget and may share deficits when spending exceeds the budget. This latter financial risk gives physician groups a strong incentive to contain spending.[8]

Economists have long concluded that medical technology is the dominant driver of health care spending growth.[9101112] Over the past half century, rapid growth in medical technology has dramatically increased treatment options for many acute and chronic conditions. For example, cardiac catheterization has expanded medicine's ability to respond to ischemic heart disease. Imaging advancements such as computed tomography (CT) and magnetic resonance imaging (MRI) have revolutionized the speed and accuracy of diagnosis. New pharmaceuticals have introduced treatment options against diseases like cancer and rheumatoid arthritis, turning them from fatal or livable diagnoses into livable chronic conditions. Advancements in surgery and image-guided interventions have broadened the scope of treatment for many conditions.

As medical technologies flourished, the appropriateness of their use became a subject of debate. In many situations, the appropriateness of an intervention is well accepted and backed by formal practice guidelines. For example, therapies for secondary prevention after a heart attack are generally both clinically effective and low-cost.[13] Colonoscopy screening for colorectal cancer is also recognized as highly beneficial, especially in older populations.[141516] In other clinical scenarios, technologies may be inappropriately used in some circumstances. For example, percutaneous coronary intervention is frequently performed for non-acute indications across U.S. hospitals, giving rise to overuse.[1718] The potential overuse of diagnostic imaging in recent decades has also come under scrutiny, such as CT or MRI for patients with low back pain without neurological symptoms or risk factors indicating a need for imaging.[1920] In the case of low back pain, reported health and functioning have not improved as spending on such imaging accrued, supporting the notion that both cost savings and clinical benefit (through avoiding unnecessary radiation exposure) could be achieved through lower use.[2122] Inappropriate imaging generated through self-referrals has been identified as an especially important problem, as physicians have increasingly owned their own diagnostic imaging equipment or facilities.[23] Even preventive care is susceptible to overuse in certain populations, such as women who receive unnecessary screening for cervical cancer after undergoing hysterectomy for benign causes.[24] Such scenarios are among the lists of low-value services produced by the American Board of Internal Medicine's "Choosing Wisely" campaign in partnership with 17 specialty societies.[2526]

For many reasons, a determination of appropriateness or "value" for any clinical service is inherently difficult. The appropriateness of any individual service depends on a variety of inputs, some of which are difficult if not impossible to measure. They include its timing in a patient's trajectory of care (removing a blood clot from the brain goes from high to low value in a matter of minutes), location of delivery (the same service in a hospital can cost much less in a clinic), and the complex clinical situation within which the treatment decision is made (patients with multiple co-morbidities). This clinical nuance extends to areas of medicine, in which the appropriateness of interventions depends on a patient's particular risk factors. Other inputs include various dimensions of patient preferences and supply-side (physician or hospital) factors, such as capacity, practice volume, and specialization, which further complicate the value equation.[27] Research suggests that in heart attack treatments, the value of any treatment strategy depends on the patient's fitness for the treatment as well as the expertise a provider has with the technology.[28] Physician expertise through the volume-outcomes relationship is also sure to play a role.[29][30] Thus, the value of any service may be quite heterogeneous across a population, with some sure to have a wider distribution around the average than others.[31]

Nevertheless, appropriateness of care has become a focus of health policy, with major legislative efforts such as the Affordable Care Act motivated, in part, by regional variations in care suggesting that perhaps one-third of U.S. health care spending is wasteful.[32][33] As the nation experiments with global budgets, understanding their effects on technology-intensive services whose appropriateness can often be unclear is important. We studied the impact of a widespread global budget initiative in Massachusetts on several such categories of services: those thought to be beneficial (colonoscopies), those thought to be overused in some cases (cardiovascular and imaging), and those thought to be driven substantially by patient preferences (orthopedics).

The Alternative Quality Contract

In 2009, Blue Cross Blue Shield of Massachusetts (BCBS) entered into global budget contracts with 7 physician organizations in Massachusetts.[34] Four additional organizations joined in 2010 with 15 groups now participating as of 2012. The "Alternative Quality Contract" (AQC) is a 5-year contract that pays organizations a global budget for the entire continuum of care for a population of enrollees in health maintenance organization plans. Enrollees designate a primary care physician (PCP) each year, and budgets are allocated to the PCP's organization. More than 1,600 PCPs and 3,200 specialists practice in organizations participating in the AQC, ranging from multi-specialty practices to large tertiary care systems. The AQC also includes pay-for-performance bonuses of up to 10% of the organization's budget, based on both inpatient and outpatient quality measures. To support AQC organizations, BCBS provides technical assistance including periodic reports that compare an organization's spending and quality to those of others.

Previous work found that the AQC was associated with a 1.9% reduction in spending and modest quality improvements in the first year.[35] This grew to a 3.3% reduction in the second year with larger improvements in quality.[36] On the whole, savings were achieved through lower prices in the first year, consistent with the organizations' focus of shifting referrals to

less expensive providers.[37] By the second year, reductions in utilization also contributed to savings.

METHODS

Study Design

The population included enrollees from January 2006 through December 2010 who were continuously enrolled for at least one calendar year. The 2009 intervention cohort consisted of 428,892 enrollees whose PCPs' organizations joined the AQC in 2009, and the 2010 cohort consisted of an additional 183,655 enrollees of organizations that joined in 2010. About 1.3 million enrollees whose PCPs did not participate in the AQC served as controls. Characteristics of the population have been described elsewhere.[38] Enrollees averaged 35 years in age, with 50% female. Average cost-sharing was about 15%. These were stable across the study period.

We used a difference-in-difference approach to characterize the treatment effect. Our primary analysis consists of the 2009 cohort, for whom the pre-intervention period was 2006-2008 and post-intervention was 2009-2010. Within this cohort, we pre-specified 2 subgroups. The "prior-risk" subgroup consisted of 4 organizations that had prior experience with risk-based contracts from BCBS (88% of the cohort), and the "no-prior-risk" subgroup consisted of 3 organizations that entered the contract without BCBS risk-contracting experience (12% of the cohort). Prior-risk organizations tended to be larger and more established in the marketplace, whereas the latter included physicians groups which were formed in preparation for entering the risk contract. We also analyzed the 2010 AQC cohort, comprising of 4 organizations without prior risk contracting, analogous to the no-prior-risk subgroup.

We identified colonoscopy, cardiovascular, imaging, and orthopedic services using Current Procedural Terminology codes and the 2010 Berenson-Eggers Type of Service (BETOS) classification system from the Centers for Medicare and Medicaid Services.[39]

Statistical Analysis

The dependent variable was spending (including patient cost sharing) in 2010 dollars or utilization. Spending was computed from claims payments made within the global budget, which reflects negotiated fee-for-service prices. Utilization was computed by counting services. For ease of interpretation, we scaled utilization data to volume per thousand enrollees.

We used a multivariate linear model at the enrollee-quarter level. We controlled for age categories, interactions between age and sex, and enrollee risk score. Risk scores were calculated by BCBS from current year demographics and diagnoses grouped by episodes of care, similar to the Hierarchical Coexisting Conditions (HCC) risk adjustment system used by the Centers for Medicare and Medicaid Services to adjust payments to Medicare Advantage plans.[40] The risk score comes from a statistical model relating current year spending to current year diagnoses and demographic information. Higher scores denote greater expected spending.

In our base model, we also included indicators for intervention status, quarter, quarter-intervention interactions, the post-intervention period, and the interaction between intervention and the post-intervention period, which produced our estimate of the policy effect. To balance the sample on observable traits, the model used propensity weights calculated from age, sex, risk score, and cost-sharing. Given the possibility of unobserved demand-side incentives, the model included indicators for each specific benefit design within BCBS HMO plans. Consistent with our prior work, the model was not logarithmic-transformed because the risk score is designed to predict dollar spending and linear models have been shown to better predict health spending than more complex functional forms.[41][42][43][44] Standard errors were clustered at the practice level.[45][46] Results are reported with 2-tailed P values.

We tested the model for differences in pre-intervention trends between treatment and control enrollees. The lack of a difference in pre-intervention trends supports our identification strategy in the difference-in-differences framework. Resulting changes in spending associated with the AQC can be explained by changes in utilization (quantity) or changes in prices. We decomposed spending results by category into price and quantity components by standardizing the prices for each service to its median price across all providers in 2006-2010. Differences in spending from repriced claims reflect differences in utilization. We further assessed whether the price effect was due to differential changes in negotiated fees or differential changes in referral patterns (referring patients to less expensive physicians or hospitals). We used our models of utilization to directly analyze the relationship between the AQC and quantity of specific services within each category.

Under a global budget, there may be incentives to upcode for increased payments, which would make patients seem sicker and thus spending adjusted for health status seem lower relative to the control group. In prior work, we explored this concern and showed that risk score changes associated with the AQC explain only a nominal share of the spending difference.[47] We used STATA software, version 11. The study was approved by Harvard Medical School.

RESULTS

Colonoscopies

Over the first two years of the AQC, the 2009 cohort saw an increase of 0.7 (p=0.04) colonoscopies per 1,000 enrollees per quarter relative to control (Figure 1), amounting to a 5.2% rise in the volume of colonoscopies (Table 1). Because colonoscopies are clinically indicated as a screening tool in patients 50 years or older, consistent with higher baseline volumes shown in Table 1, we repeated our analysis in this subpopulation. Estimates show a statistically insignificant increase of 1.9 (p=0.13) colonoscopies per 1,000 enrollees per quarter, or about 4.7%, relative to control. In the 2010 cohort, no changes in colonoscopy utilization associated with the AQC were found in its first year (−0.3, p=0.48).

Spending on colonoscopies increased by 5.7% over the first 2 years in the 2009 cohort ($0.54 per enrollee per quarter). As Table 2 shows, this increase was statistically significant in year-1 (p=0.005) but not significant in year-2 (p=0.22). The 2010 cohort saw no changes

in colonoscopy spending relative to control in its first year ($0.10, p=0.88). Overall, changes in colonoscopy spending were driven by the changes in volume, as we did not find significant changes in colonoscopy prices or location of care associated with the AQC.

Cardiovascular Services

Spending on cardiovascular services was higher in the post-AQC period relative to pre-AQC for both intervention and control subjects, but the difference was smaller in AQC subjects. The 2009 AQC cohort spent on average 7.4% less (−$1.47 per member per quarter, p=0.02) relative to control. By year, these savings were 6.8% (p=0.001) in year-1 and 7.7% (p=0.04) in year-2 (Table 2). The greater difference in year 2 was unrelated to the removal of the 2010 cohort from the control group in year 2 (the 2010 cohort belonged to the control group in year 1). Although selection into the AQC was non-random, an interaction of the secular trend with the AQC indicator demonstrated no significant spending trend differences between AQC and non-AQC groups prior to the intervention. This suggests that the AQC effect is not explained by differential underlying trends in spending between the groups. The 2010 cohort spent 11.1% less (−$2.54, p=0.004) on cardiovascular services relative to control its first year.

For the 2009 cohort, about one-third of the savings was explained by lower utilization of services, while two-thirds by lower prices. In our decomposition of the price effect, we found no differences in price trends between AQC and non-AQC providers. Instead, spending reductions due to prices were explained by patients receiving care in outpatient facilities with less expensive fees, consistent with our prior work.[47] Direct analyses of cardiovascular services (coronary artery bypass grafts, aneurysm repairs, endarterectomies, angioplasties, and pacemaker insertions) are reported in Table 1 and plotted in Figure 2. Results suggest that a decrease in angioplasty volume likely drove the differences in spending. Other cardiovascular services did not undergo changes in volume, consistent with Figure 2.

Spending reductions were larger in the no-prior-risk subgroup, which had an average reduction of 27.2% (−$6.06, p<0.001) in cardiovascular spending relative to control in the first two years. In contrast, the prior-risk subgroup experienced smaller reductions of 4.1% (−$0.85, p=0.17). Sensitivity analyses supported our main results.

Imaging

The 2009 AQC cohort spent less on imaging relative to control in the first 2 years. The AQC was associated with a 6.1% decrease (−$5.67 per member per quarter, p=0.001) on imaging spending. The year-1 reduction was 4.7% (−$4.40, p<0.001) and year-2 was 7.4% (−$6.86, p=0.001) (Table 2). There were no significant differences in pre-intervention spending trends on imaging between the AQC and non-AQC groups.

Spending reductions were found in the outpatient setting rather than the hospital. In addition, outpatient facility spending (the portion of the fee that goes to the facility) accounted for almost all of the savings, suggesting that imaging services were referred to less expensive settings (such as non-hospital based), giving rise to a large price effect that drove the findings. Consistent with this and Figure 3, direct analyses of utilization found no

statistically significant decreases in the volume of CTs, MRIs, ultrasounds/echocardiograms, or imaging procedures (Table 1). Standard imaging, mostly comprised of X-rays, actually saw an increase in volume of 2.5% (p=0.045) over the first 2 years (6.19 per 1000 enrollees per quarter). This further implies a substantial price effect that was able to overcome some volume increases to produce savings.

Similar to cardiovascular services, savings were larger in the no-prior-risk subgroup. In this subgroup, the AQC was associated with a reduction of 11.9% on imaging spending (−$10.54 per enrollee per quarter, p=0.02) after two years. The 2010 cohort, all no-prior-risk groups, saw a similar reduction of 11.0% (−$10.90, p<0.001) in its first year. In the prior-risk subgroup, savings were smaller. The average reduction after 2 years was 4.7% (−$4.51, p<0.001).

Orthopedics

The AQC was not associated with any changes in spending on orthopedic services, nor were there any associated changes in utilization. The average 2-year effect of the AQC on total orthopedic spending was an insignificant 0.9% ($0.04 per enrollee per quarter, p=0.83). Neither year-1 nor year-2 saw significant effects (Table 2). This was consistent across prior-risk and no-prior-risk groups in the 2009 cohort. In addition, neither facility nor non-facility components of orthopedic spending were affected. The 2010 cohort saw a large though statistically insignificant reduction in orthopedic spending of 12.7% (−$0.70, p=0.09).

Consistent with Figure 4, direct analyses of utilization demonstrated that the AQC was not associated with changes in the volume of hip or knee replacements (Table 1).

DISCUSSION

The AQC was associated with increased spending on colonoscopies due to increased volume. It was associated with reductions in spending on cardiovascular and imaging services, but not on orthopedic services. Reductions in spending were larger in year-2 than in year-1, and were explained both by decreased utilization and lower prices achieved through referring patients to less expensive providers. Consistent with qualitative evidence gathered from organizational leaders, shifting referral patterns to less expensive providers was the low-hanging fruit by which organizations in the AQC aimed to achieve savings in the early years.[48] However, among both cardiovascular and imaging services, about one-third of the reductions in spending, concentrated in year-2 and in groups who entered the contract from fee-for-service, was explained by lower volume. This suggests that even in the early years, organizations were able to achieve reductions in volume in some services.

Under bundled and global payment systems, there can be a tradeoff between controlling spending and supporting innovation. Innovations in medicine often lead to technologies that are high cost, such as new devices or diagnostic modalities, some of which may be high or low value depending on the clinical situation. In cases such as angioplasty and imaging, extending these innovations to patients in certain situations, such as a lack of clinical indications for use or purely elective use, have been questioned on the grounds of appropriateness. One way that systems operating under global payments have to reduce

spending while maintaining quality is to lower spending on low-value services. Thus, a crucial question surrounding our findings is whether the foregone spending was for high or low value services. While our analysis of claims data does not include any conclusive assessments of the effect on value, several pieces of supporting evidence that suggest the AQC did not adversely affect the quality of care and may have successfully targeted services that may be overused in some situations.

First, in other research we found that the AQC did not negatively affect the pay-for-performance measures of quality used in the AQC program. In contrast, we found that the AQC was associated with improvements in quality in chronic care management, adult preventive care, as well as pediatric care.[49] Second, the AQC was associated with increased use of colonoscopies, a service generally supported by practice guidelines. Third, our findings on imaging (Figure 3) are consistent with broader trends in the slow down of imaging, which is posited to be driven largely by reductions in use in lower-value situations.[50] Finally, our cardiology results are also consistent with a focus on value. Our results merely provide suggestions, rather than conclusive evidence, about value. We must emphasizing that value depends on the clinical situation.

Our study has several limitations. First, the population was drawn from a large commercially insured HMO population in Massachusetts, so results may not generalize to Medicare or enrollees in other types of health plans. Second, we did not observe details of each AQC contract or provider risk contracting with other payers, which have become more prevalent in 2010. Third, using administrative data, we cannot assess the appropriateness of any specific instance of utilization. Fourth, our service categories do not capture the broad array of medical technologies that patients and doctors use. We merely explored a small mixture of services. In addition, beliefs about appropriateness may be in flux; for example, recent practice guidelines for colonoscopies have also called into question its presumed high value.[51] Finally, we caution that multiple comparisons within categories of services may produce spurious results.

Slowing the growth of health care spending without hurting quality is a central goal.[52][53] As payment reform and ACOs gain momentum across the country, and physician organizations adapt to a more constrained environment, avoiding the blunting of technological innovations in medicine is an important priority.[54] Different medical technologies will likely be affected differently by global budgets. This will depend on providers' and patients' treatment choices, but it will also depend on the ability of device makers and manufacturers of tools, scanners, and other technologies to focus on innovations that are high value. As lessons from early global or bundled payment systems amass,[55][56] an understanding of the mechanisms by which spending is reduced will be as important to policymakers as that it can be reduced at all.

Acknowledgments

References

1. Chernew ME, Baicker K, Hsu J. The specter of financial armageddon—health care and federal debt in the United States. *N Engl J Med.*. Apr 1; 2010 362(13):1166–8. [PubMed: 20237338]

2. Antos JR, Pauly MV, Wilensky GR. Bending the Cost Curve through Market-Based Incentives. N Engl J Med. Sep 6; 2012 367(10):954–8. [PubMed: 22852881]

3. Emanuel E, Tanden N, Altman S, et al. A Systemic Approach to Containing Health Care Spending. N Engl J Med. Sep 6; 2012 367(10):949–54. [PubMed: 22852883]

4. Aaron HJ. The Central Question for Health Policy in Deficit Reduction. N Engl J Med. 2011; 365:1655–1657. [PubMed: 21995336]

5. Fisher ES, McClellan MB, Safran DG. Building the Path to Accountable Care. N Engl J Med. 2011; 365:2445–2447. [PubMed: 22204720]

6. Frakt AB, Mayes R. Beyond capitation: how new payment experiments seek to find the 'sweet spot' in amount of risk providers and payers bear. Health Aff (Millwood). Sep; 2012 31(9):1951–8. [PubMed: 22949443]

7. Meyer H. Many Accountable Care Organizations Are Now Up And Running, If Not Off To The Races. Health Aff (Millwood). Nov; 2012 31(11):2363–7. [PubMed: 23129665]

8. Engleberg Center for Health Care Reform. Bending the Curve: Effective Steps to Address Long-Term Health Care Spending Growth. The Brookings Institution; Aug. 2009 http://www.brookings.edu/reports/2009/0901_btc.aspx

9. Newhouse JP. Medical care costs: how much welfare loss? J Econ Perspect. 1992; 6(Summer)(3):3–21. [PubMed: 10128078]

10. Garber AM, Skinner J. Is American Health Care Uniquely Inefficient? J Econ Perspect. 2008; 22(Fall)(4):27–50. [PubMed: 19305645]

11. Chernew, ME.; Newhouse, JP. Health Care Spending Growth. In: Pauly, Mark V.; McGuire, Thomas G.; Barros, Pedro P., editors. Handbook of Health Economics. Vol. 2. Elsevier Science; North Holland: 2012. p. 1-43.

12. Cutler, DM. Your Money or Your Life: Strong Medicine for America's Health Care System. 2005.

13. Smith SC Jr, Benjamin EJ, Bonow RO, et al. AHA/ACCF secondary prevention and risk reduction therapy for patients with coronary and other atherosclerotic vascular disease: 2011 update: A guideline from the American Heart Association and American College of Cardiology Foundation. Circulation. 2011; 124:2458–2473. [PubMed: 22052934]

14. Sonnenberg A, Delco F, Inadomi JM. Cost-effectiveness of colonoscopy in screening for colorectal cancer. Ann Intern Med. 2000; 133:573–84. [PubMed: 11033584]

15. Frazier AL, Colditz GA, Fuchs CS, et al. Cost-effectiveness of screening for colorectal cancer in the general population. JAMA. 2000; 284:1954–61. [PubMed: 11035892]

16. Winawer S, Fletcher R, Rex D, et al. Colorectal cancer screening and surveillance: clinical guidelines and rationale—update based on new evidence. Gastroenterology. 2003; 124:544–60. [PubMed: 12557158]

17. Chan PS, Patel MR, Klein LW, et al. Appropriateness of percutaneous coronary intervention. JAMA. Jul 6; 2011 306(1):53–61. [PubMed: 21730241]

18. Bradley SM, Maynard C, Bryson CL. Appropriateness of percutaneous coronary interventions in Washington State. Circ Cardiovasc Qual Outcomes. Jul 1; 2012 5(4):445–53. [PubMed: 22570356]

19. Chou R, Deyo RA, Jarvik JG. Appropriate use of lumbar imaging for evaluation of low back pain. *Radiol Clin North Am.*. Jul; 2012 50(4):569–85. [PubMed: 22643385]

20. Webster BS, Courtney TK, Huang YH, Matz S, Christiani DC. Physicians' initial management of acute low back pain versus evidence-based guidelines. Influence of sciatica. *J Gen Intern Med.*. 2005:201132–5.

21. Martin BI, Deyo RA, Mirza SK, Turner JA, Comstock BA, Hollingworth W, et al. Expenditures and health status among adults with back and neck problems. JAMA. 2008; 299:656–64. [PubMed: 18270354]

22. Chou R, Qaseem A, Owens DK, et al. Diagnostic imaging for low back pain: advice for high-value health care from the American College of Physicians. *Ann Intern Med.*. Feb 1; 2011 154(3):181–9. [PubMed: 21282698]

23. Paxton BE, Lungren MP, Srinivasan RC, et al. Physician self-referral of lumbar spine MRI with comparative analysis of negative study rates as a marker of utilization appropriateness. *Am J Roentgenol.*. Jun; 2012 198(6):1375–9. [PubMed: 22623551]

24. Cervical Cancer Screening Among Women by Hysterectomy Status and Among Women Aged 65 Years — United States, 2000-2010. Morbidity and Mortality Weekly Report (MMWR). Jan 4; 2013 61(51):1043–1047.

25. Choosing Wisely. An initiative of the American Board of Internal Medicine. www.choosingwisely.org

26. Rao VM, Levin DC. The overuse of diagnostic imaging and the Choosing Wisely initiative. *Ann Intern Med.*. Oct 16; 2012 157(8):574–6. [PubMed: 22928172]

27. Chandra, A.; Cutler, D.; Song, Z. Who Ordered That? The Economics of Treatment Choices in Medical Care. In: Pauly, MV.; McGuire, TG.; Barros, PP., editors. Handbook of Health Economics. Vol. 2. Elsevier Science; North Holland: 2012. p. 397-432.

28. Chandra A, Staiger DO. Productivity spillovers in health care: Evidence from the treatment of heart attacks. Journal of Political Economy. 2007; 115(1):103–140. [PubMed: 18418468]

29. Birkmeyer JD, Siewers AE, Finlayson EV, et al. Hospital volume and surgical mortality in the United States. New England Journal of Medicine. 2002; 346:1128–1137. [PubMed: 11948273]

30. Van Heek NT, Kuhlmann KF, Scholten RJ, de Castro SM, Busch OR, van Gulik TM, et al. Hospital volume and mortality after pancreatic resection: A systematic review and an evaluation of intervention in the Netherlands. Annals of Surgery. 2005; 242(6):781–788. [PubMed: 16327488]

31. Chandra, Amitabh; Skinner, Jonathan. Technology Growth and Expenditure Growth in Health Care. Journal of Economic Literature. 2012; 50(3):645–80.

32. Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder EL. The implications of regional variations in Medicare spending. Part 1: the content, quality, and accessibility of care. Ann Intern Med. Feb 18; 2003 138(4):273–87. [PubMed: 12585825]

33. Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder EL. The implications of regional variations in Medicare spending. Part 2: health outcomes and satisfaction with care. Ann Intern Med. Feb 18; 2003 138(4):288–98. [PubMed: 12585826]

34. Chernew ME, Mechanic RE, Landon BE, Safran DG. Private-Payer Innovation In Massachusetts: The 'Alternative Quality Contract.'. Health Aff (Millwood). 2011; 30(1):51–61. [PubMed: 21209437]

35. Song Z, Safran DG, Landon BE, He Y, Ellis RP, Mechanic RE, Day MP, Chernew ME. Health care spending and quality in year 1 of the alternative quality contract. N Engl J Med. Sep 8; 2011 365(10):909–18. [PubMed: 21751900]

36. Song Z, Safran DG, Landon BE, Landrum MB, He Y, Mechanic RE, Day MP, Chernew ME. The 'alternative quality contract,' based on a global budget, lowered medical spending and improved quality. Health Aff (Millwood). Aug; 2012 31(8):1885–94. [PubMed: 22786651]

37. Mechanic RE, Santos P, Landon BE, Chernew ME. Medical group responses to global payment: early lessons from the 'Alternative Quality Contract' in Massachusetts. Health Aff (Millwood). Sep; 2011 30(9):1734–42. [PubMed: 21900665]

38. Song Z, Safran DG, Landon BE, Landrum MB, He Y, Mechanic RE, Day MP, Chernew ME. The 'alternative quality contract,' based on a global budget, lowered medical spending and improved quality. Health Aff (Millwood). Aug; 2012 31(8):1885–94. [PubMed: 22786651]

39. Centers for Medicare and Medicaid Services. Berenson-Eggers Type of Service. 2010. Available at: https://www.cms.gov/HCPCSReleaseCodeSets/20_BETOS.asp

40. Pope GC, Kautter J, Ellis RP, Ash AS, Ayanian JZ, Iezzoni LI, Ingber MJ, Levy JM, Robst J. Risk Adjustment of Medicare Capitation Payments Using the CMS-HCC model. Health Care Financing Review. 2004; 25(4):119–41. [PubMed: 15493448]

41. Manning WG, Basu A, Mullahy J. Generalized modeling approaches to risk adjustment of skewed outcomes data. J Health Econ. 2005; 24(3):465–88. [PubMed: 15811539]

42. Zaslavsky AM, Buntin MB. Too much ado about two-part models and transformation? Comparing methods of modeling Medicare expenditures. J Health Econ. 2004; 23:525–42. [PubMed: 15120469]

43. Ai C, Norton EC. Interaction terms in logit and probit models. Economic Letters. 2003; 80:123–9.

44. Ellis RP, McGuire TG. Predictability and predictiveness in health care spending. J Health Econ. 2007; 26(1):25–48. [PubMed: 16908083]

45. White H. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica. 1980; 48(4):817–30.

46. Huber, PJ. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. University of California Press; Berkeley: 1967. The behavior of maximum likelihood estimates under non-standard conditions; p. 221-33.

47. Song Z, Safran DG, Landon BE, He Y, Ellis RP, Mechanic RE, Day MP, Chernew ME. Health Care Spending and Quality in Year 1 of the Alternative Quality Contract. N Engl J Med. 2011; 365:909–18. [PubMed: 21751900]

48. Mechanic RE, Santos P, Landon BE, Chernew ME. Medical group responses to global payment: early lessons from the 'Alternative Quality Contract' in Massachusetts. Health Aff (Millwood). Sep; 2011 30(9):1734–42. [PubMed: 21900665]

49. Song Z, Safran DG, Landon BE, Landrum MB, He Y, Ellis RP, Mechanic RE, Day MP, Chernew ME. The 'Alternative Quality Contract' Based On A Global Budget, Lowered Medical Spending And Improved Quality. Health Aff(Milwood). 2012; 31(8):1885–1894.

50. Lee DW, Levy F. The sharp slowdown in growth of medical imaging: an early analysis suggests combination of policies was the cause. Health Aff (Millwood). Aug; 2012 31(8):1876–84. [PubMed: 22842655]

51. Austin GL, Fennimore B, Ahnen DJ. Can Colonoscopy Remain Cost-Effective for Colorectal Cancer Screening? The Impact of Practice Patterns and the Will Rogers Phenomenon on Costs. Am J Gastroenterol. Dec 4.2012

52. Berwick DM. Making Good on ACOs' Promise — The Final Rule for the Medicare Shared Savings Program. N Engl J Med. 2011; 365:1753–1756. [PubMed: 22013899]

53. McClellan M, McKethan AN, Lewis JL, Roski J, Fisher ES. A national strategy to put accountable care into practice. Health Aff (Millwood). 2010; 29(5):982–90. [PubMed: 20439895]

54. Altman SH. The Lessons Of Medicare's Prospective Payment System Show That The Bundled Payment Program Faces Challenges. Health Aff (Millwood). Sep; 2012 31(9):1923–30. [PubMed: 22949439]

55. Weissman JS, Bailit M, D'Andrea G, Rosenthal MB. The design and application of shared savings programs: lessons from early adopters. Health Aff (Millwood). Sep; 2012 31(9):1959–68. [PubMed: 22949444]

56. Fisher ES, Shortell SM, Kreindler SA, Van Citters AD, Larson BK. A framework for evaluating the formation, implementation, and performance of accountable care organizations. Health Aff (Millwood). Nov; 2012 31(11):2368–78. [PubMed: 23129666]
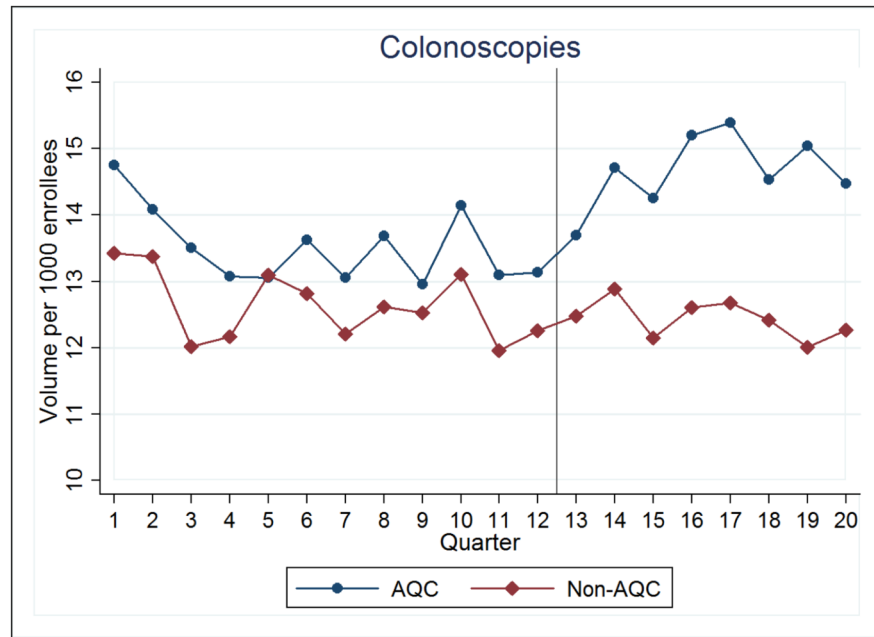
Figure 1. Colonoscopy Utilization

For all figures, AQC denotes the 2009 intervention cohort, and Non-AQC denotes the control. The x-axis represents 2006-2010 in quarters, with the vertical line placed at the start of 2009 when the AQC was implemented.
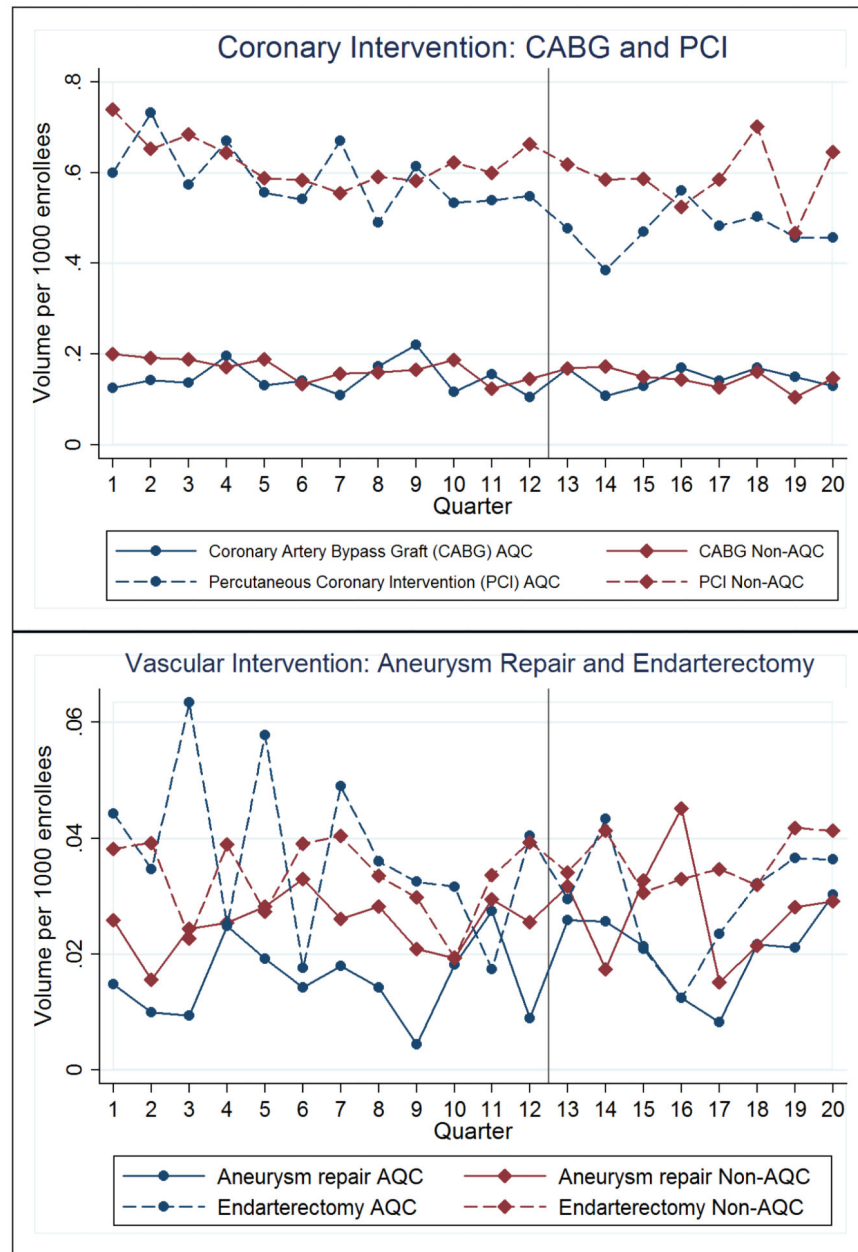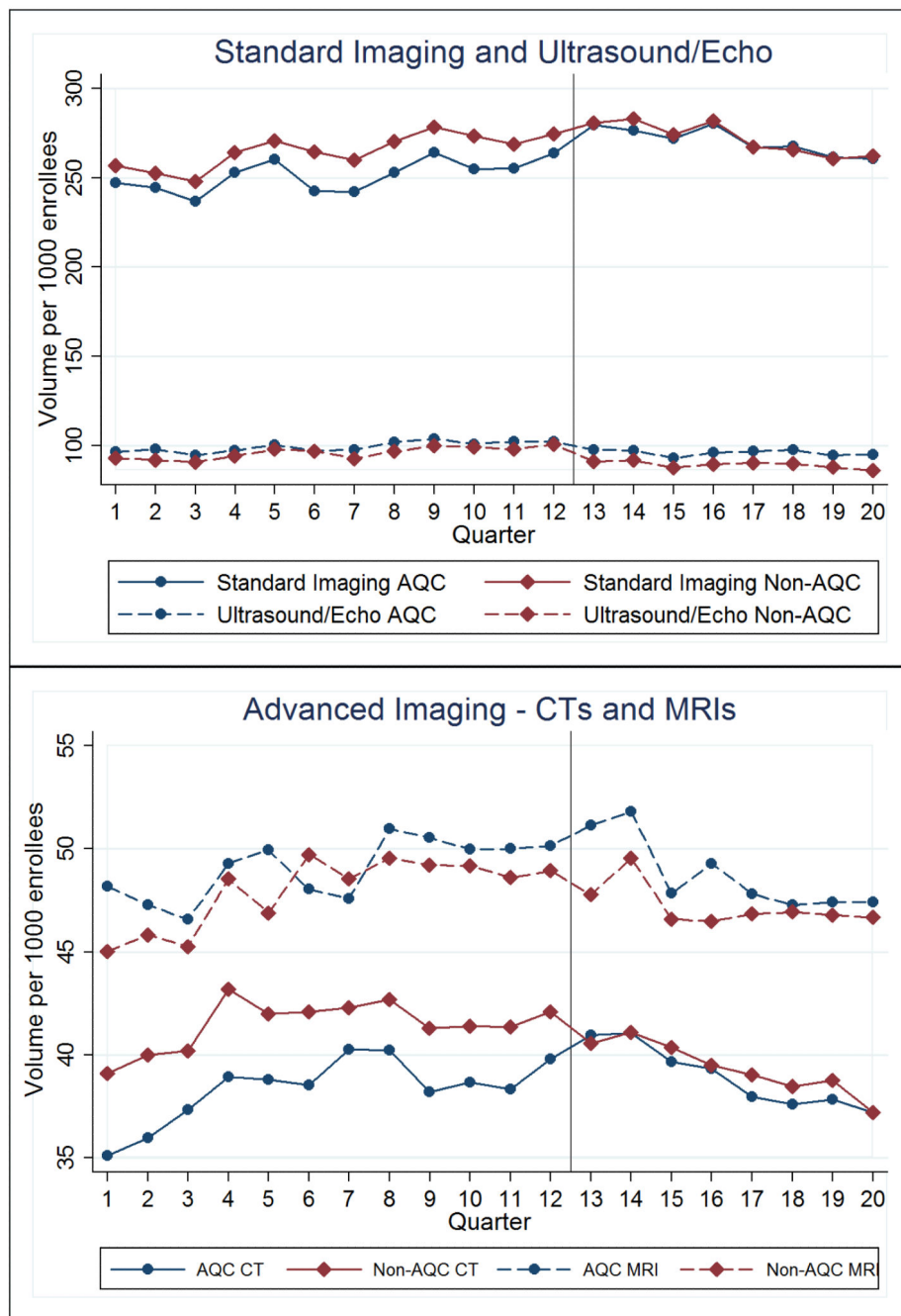
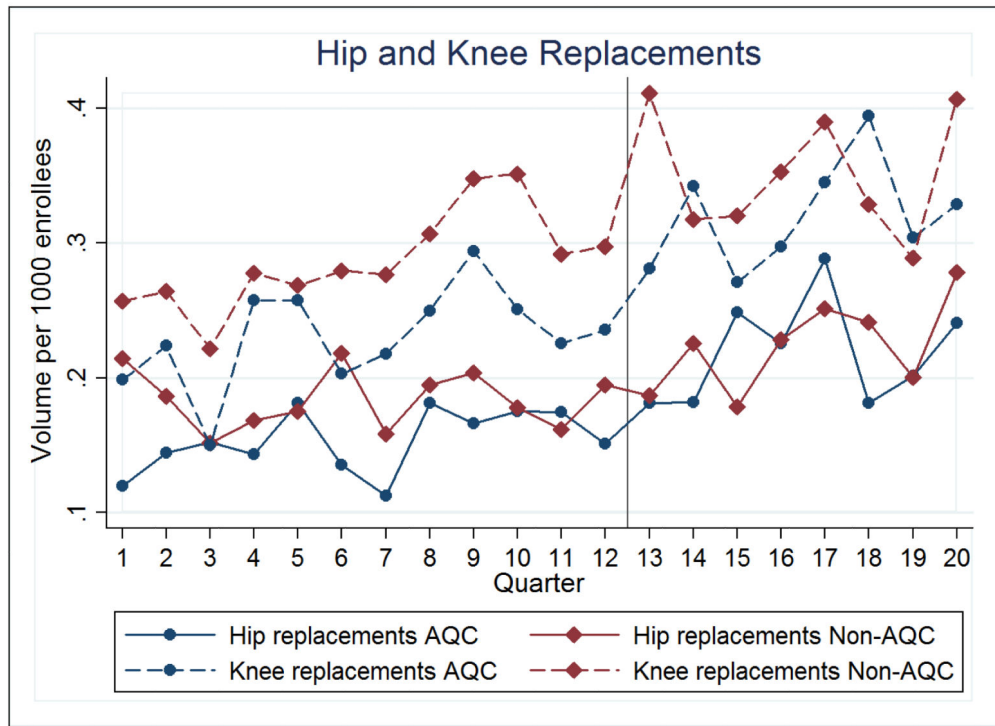**Figure 2. Cardiovascular Utilization**

Figure 3. Imaging Utilization

Figure 4. Orthopedic Utilization

Table 1

Changes in Utilization in Treatment and Control Groups (volume per 1000 enrollees per quarter)

Category	Treatment Enrollees (AQC) (N=428,892)			Control Enrollees (Non-AQC) (N=1,339,798)			Between-Group Difference Average 2-year effect (2009-10)		
	Pre-AQC (2006-08)	Post-AQC (2009-10)	Change	Pre-AQC (2006-08)	Post-AQC (2009-10)	Change	Unadjusted	Adjusted	P-value
Colonoscopy									
All enrollees	13.53	14.60	1.07	12.62	12.36	−0.26	1.33	0.70	0.04
Over 50	40.76	40.95	0.19	35.09	33.11	−1.98	2.17	1.90	0.13
Cardiovascular									
CABG	0.14	0.15	0.01	0.17	0.15	−0.02	0.03	0.01	0.60
Aneurysm repair	0.01	0.02	0.01	0.02	0.03	0.01	0.00	0.00	0.80
Endarterectomy	0.04	0.03	−0.01	0.03	0.04	0.01	−0.02	−0.01	0.07
Angioplasty	0.58	0.49	−0.09	0.62	0.60	−0.02	−0.07	−0.12	0.02
Pacemaker	0.14	0.15	0.01	0.18	0.19	0.01	0.00	−0.02	0.41
Other	3.71	2.85	−0.86	4.26	3.00	−1.26	0.40	0.15	0.46
Imaging									
Standard imaging	251.99	269.85	17.86	265.21	271.10	5.89	11.97	6.19	0.045
CT	38.23	39.21	0.98	41.30	39.59	−1.71	2.69	1.16	0.07
MRI	49.18	48.50	−0.68	48.00	46.95	−1.05	0.37	−1.06	0.30
Ultrasound/Echo	99.56	95.80	−3.76	96.22	89.00	−7.22	3.46	0.47	0.72
Imaging procedures	13.26	15.67	2.41	14.28	16.03	1.75	0.66	0.07	0.87
Orthopedics									
Hip replacement	0.15	0.22	0.07	0.18	0.22	0.04	0.03	0.02	0.29
Knee replacement	0.23	0.31	0.08	0.29	0.35	0.06	0.02	0.01	0.63

Table 2

Changes in Spending in Treatment and Control Groups ($ per enrollee per quarter)

Category	Treatment Enrollees (AQC) (N=428,892)			Control Enrollees (Non-AQC) (N=1,339,798)			Between-Group Difference							
							Average 2-year effect			Year-1 effect		Year-2 effect	
	Pre-AQC	Post-AQC	Change	Pre AQC	Post AQC	Change	Unadj.	Adj.	P		P		P	
Colonoscopy														
All enrollees	9.42	11.64	2.22	9.39	10.65	1.26	0.96	0.54	0.16	0.55	0.005	0.56	0.22
Over 50 yrs	27.75	31.73	3.98	25.39	27.78	2.39	1.59	1.37	0.24	1.58	0.01	1.24	0.38
Cardiovascular	20.03	22.17	2.14	22.40	25.40	3.00	−0.86	−1.47	0.02	−1.37	0.001	−1.54	0.04
Imaging	93.10	106.54	13.44	102.31	118.80	16.49	−3.05	−5.67	0.001	−4.40	<0.001	−6.86	0.001
Orthopedics	4.12	5.68	1.56	4.88	6.30	1.42	0.14	0.04	0.83	−0.01	0.95	0.07	0.77