Variation In Accountable Care Organization Spending And Sensitivity To Risk Adjustment: Implications For Benchmarking

Sherri Rose,

Department of Health Care Policy, Harvard Medical School, Boston, Massachusetts

Alan M. Zaslavsky, and

Department of Health Care Policy, Harvard Medical School, Boston, Massachusetts

J. Michael McWilliams

Department of Health Care Policy, Harvard Medical School, 180 Longwood Ave, Boston, Massachusetts

Abstract

Spending targets (or benchmarks) for accountable care organizations (ACOs) participating in the Medicare Shared Savings Program must be carefully set to encourage program participation while achieving fiscal goals and minimizing unintended consequences, such as penalizing ACOs for serving sicker patients. Recently proposed regulatory changes include measures to make benchmarks more similar for ACOs in the same area with different historical spending levels. We found that ACOs vary widely in how their spending levels compare with other local providers after standard case-mix adjustments. Additionally adjusting for survey measures of patient health meaningfully reduced the variation in differences between ACO spending and local average fee-for-service spending, but substantial variation remained, suggesting that differences in care efficiency between ACOs and local non-ACO providers vary widely. Accordingly, measures to equilibrate benchmarks between high- and low-spending ACOs--such as setting benchmarks to risk-adjusted average fee-for-service spending in an area--should be implemented gradually to maintain participation by ACOs with high spending. Use of survey information also could help mitigate perverse incentives for risk selection and upcoding and limit unintended consequences of new benchmarking methodologies for ACOs serving sicker patients.

With more than 420 participating organizations and growing, the Medicare accountable care organization (ACO) programs (the Medicare Shared Savings Program and the Pioneer model) are at the leading edge of federal efforts to transition from paying providers by fee-for-service to alternative models intended to reward more efficient care, that is, care that produces better health outcomes at lower costs.

In the Medicare Shared Savings Program--the largest of Medicare's ACO programs--the Centers for Medicare and Medicaid Services (CMS) sets a spending target or financial benchmark for each ACO. For the initial three-year contract period, this benchmark is set at the average level of Medicare spending for patients served by the ACO during its preceding three-year baseline period, updated for each contract year based upon national fee-for-service Medicare spending growth since the baseline period. If an ACO keeps spending for an attributed population of Medicare patients sufficiently below its benchmark, it shares in

the savings with Medicare, with the proportion of savings received by the ACO determined by how well it performs on a set of quality measures.[1]

ACO Benchmark Setting--An Evolving Policy

The process by which the benchmark is set is critically important to the viability of CMS's voluntary ACO programs. This is because CMS, in setting benchmarks, must establish incentives that are sufficiently strong to encourage organizations to participate as ACOs and generate savings while also achieving fiscal goals and minimizing unintended consequences, such as penalizing ACOs for serving sicker patients. Under current rules, an ACO's spending level during its most recent three-year period serves as the basis for resetting ("rebasing") its benchmark for it next three-year contract period. This method diminishes incentives for ACOs to reduce spending because they are penalized with lower benchmarks (instead of being rewarded) in the next contract period.[2] Recently, CMS proposed to transition benchmarks for ACOs closer to the level of risk-adjusted fee-for-service spending in their service area. [3] Specifically, the proposed rule would blend an ACO's rebased historical average spending after its initial contract period with average risk-adjusted spending in its region, giving increasing weight to the regional component over successive contract periods (35% in the second contract period and 70% in the third).

Because regional spending is determined by all providers in the area, the proposed rules would weaken the link between an ACO's benchmark and its prior savings, strengthening incentives for participating providers to generate savings. Because savings may be easier to achieve for ACOs that have high baseline spending and more opportunities to reduce wasteful care,[4] the proposed change also addresses concerns that the current benchmarking methodology disproportionately rewards ACOs that are more inefficient at baseline and does not reward ACOs for maintaining levels of spending below a benchmark. [5]

Challenges Ahead

The proposed measures to induce convergence in benchmarks between ACOs with baseline spending above versus below local average fee-for-service spending face three major challenges. First, excessively rapid convergence could cause ACOs with higher baseline spending to leave the program as benchmarks fall below their reach. Exiting the program would be particularly likely among those ACOs participating in two-sided risk contracts, under which they receive bonuses if their spending is below benchmarks but also incur penalties if their spending exceeds benchmarks. Incentives to exit would be weaker under one-sided contracts, in which ACOs receive bonuses if their spending performance is sufficiently below the benchmark, but do not face penalties if spending exceeds it (shared savings only). Similarly, rapid convergence could also discourage ACOs with high spending from entering two-sided contracts. Currently, ACOs in the Shared Savings Program are not required to assume downside risk during their first two contract periods, but higher shared savings rates are available to those that do. Substantial convergence of benchmarks toward a regional average before ACOs with high baseline spending are able to lower their spending could also lead to particularly generous benchmarks for ACOs with low baseline spending

(since benchmarks would reflect the average of high-and low-spending providers in the area). Such benchmarks could reward already efficient ACOs handsomely for maintaining the status quo, thereby partially or fully offsetting any savings to Medicare from the ACO Programs.

The amount of variation in ACO spending levels relative to market averages could inform expectations about how quickly convergence should occur; greater variation would suggest a need for slower convergence to allow sufficient time for ACOs with higher spending to become more efficient. To our knowledge, however, this variation has not yet been quantified.

Second, methodologies that move away from using an ACO's historical spending as the basis for its benchmark must rely more heavily on risk adjustment to ensure that ACOs are not unfairly penalized if they have higher spending because they serve sicker beneficiaries and not because they are less efficient. Similarly, the proposed use of regional spending averages as benchmarks could inappropriately reward ACOs for serving healthier patients if risk adjustment is inadequate, thereby establishing incentives for ACOs to attract healthier patients rather than become more efficient. Currently, an ACO's benchmark is adjusted for changes in the composition of the population attributed to the ACO based on the Hierarchical Condition Category (HCC) model, which predicts spending in a year of interest based on diagnosis codes in the prior year of claims. The regional component of proposed benchmarking methodology also would be adjusted to reflect an ACO's case mix using the HCC model. The extent to which patients vary across ACOs in health-related or sociodemographic factors that are neither included in the HCC model nor reflected in local fee-for-service use is unknown.

Third, no matter how ACO benchmarks are set, approaches to risk-adjusting ACO benchmarks must contend with the fact that providers supply the diagnoses used for risk adjustment. The proposed rule calls for greater scrutiny of coding practices by ACOs but does not describe specific measures to address changes in coding practices. Under current methods of risk-adjusting benchmarks, ACOs have incentives to code more intensively (or "upcode") as a means to increase HCC scores and therefore their spending benchmarks. Such changes in coding practices have been observed in Medicare Advantage as a consequence of analogous incentives introduced by the implementation of the HCC model. [6] Patient-reported measures of health may help distinguish true changes in case-mix from changes in coding practices and could be used to enhance existing risk adjustment methods if sufficiently predictive of spending.

In this study we analyzed survey and Medicare claims data to estimate differences between each ACO's spending and the spending of other, non-ACO providers in a geographic area, after standard case-mix adjustment. We found wide variation in these differences. Making additional adjustments for survey measures of patient health meaningfully reduced the variation in differences between ACO spending and local average fee-for-service spending, but substantial variation remained. These results suggest that differences in care efficiency between ACOs and local non-ACO providers vary widely and have important implications for ACO benchmarking policy.

Study Data And Methods

Overview Of Analysis

We used data from the Consumer Assessment of Healthcare Providers and Systems (CAHPS) survey of fee-for-service Medicare beneficiaries and the survey participants' linked Medicare claims from 2011 to 2012. We conducted two analyses to inform consideration of changes to ACO benchmarking policy. First, we quantified the amount of variation across ACOs in the difference between an ACO's spending level and local average spending for beneficiaries not attributed to an ACO, adjusting for beneficiaries' demographic characteristics and HCC score. Second, we assessed the impact of adjustment for health characteristics measured in the CAHPS survey on this variation in ACO spending deviations, with and without concurrent adjustment for HCC score.

Data Source And Study Population

We analyzed data from the 2011–12 CAHPS survey participants and their linked Medicare claims spanning the survey year and prior year. The fee-for-service CAHPS survey is administered annually to a random sample of community-dwelling fee-for-service beneficiaries.[7] Enrollees in both Medicare and Medicaid were surveyed in 2012 but not in 2011. We limited our sample to participants continuously enrolled in both parts A and B of traditional Medicare in the prior year and in the survey year (while alive for those who had died). We also restricted the sample to participants with at least one qualifying evaluation and management service during the survey year to support the attribution of patients to an ACO or to a non-ACO provider.

Following Medicare Shared Savings Program assignment rules and previously described methods,[1,4,8] we attributed each participant in each survey year to the ACO or non-ACO taxpayer identification number accounting for the most allowed charges for qualifying outpatient evaluation and management services provided to the participant. We excluded from consideration evaluation and management services in nursing facilities, to maintain consistency across ACOs; some ACOs include postacute and long-term care providers in their contracts, while others do not. We included the 250 Shared Savings Program or Pioneer ACOs entering contracts in 2012 or 2013 in our analysis, using CMS data to define ACOs.

To test the sensitivity of our estimates to differences between CAHPS respondents and the full community-dwelling population of fee-for-service Medicare beneficiaries (for example, because of survey nonresponse), we also analyzed Medicare claims for a random 20 percent sample of fee-for-service beneficiaries in 2011–12. We applied the same inclusion criteria to this sample, using a validated claims-based algorithm to exclude nursing home residents.[9]

Patient Characteristics

From Medicare enrollment data, we assessed age, sex, disability as the original reason for Medicare entitlement, Medicaid coverage, and presence of end-stage renal disease. From diagnoses in claims in the year prior to the survey, we determined each participant's HCC risk score. We also assessed beneficiaries' history of twenty-seven chronic conditions from the Chronic Condition Data Warehouse (CCW), which has drawn on claims data since 1999

to describe beneficiaries' accumulated burden of chronic disease.[10] Specifically, for each condition, we created an indicator of ever having the condition prior to the start of the survey year.

From the CAHPS survey, we assessed the following variables directly related to health: general health status, mental health status, difficulties with activities of daily living, smoking status, and requiring help to complete the survey (which has been shown to be strongly related to cognitive impairment).[11] We also assessed the following social and demographic factors: race and ethnicity, educational attainment, living alone, and sources of supplemental and prescription drug coverage (Medigap, employer-sponsored insurance, Veterans Affairs or other military retiree benefits, and prescription drug plans). (See online Appendix Exhibit A1 for relevant CAHPS survey questions. Item nonresponse rates were generally low and are presented in Appendix Exhibit A2.)[12] We handled item nonresponse analytically by including a missing category for each characteristic in statistical models.

Spending

The primary outcome variable was total Medicare spending in the concurrent (survey) year, as determined from the Medicare Provider Analysis and Review File (to assess payments to facilities for inpatient and postacute care), the Carrier (physician/supplier) File, the Outpatient File, and the cost and use segment of the Master Beneficiary Summary File (to assess spending on hospice care, home health care, and durable medical equipment). Because the CAHPS survey is administered early in the calendar year (more than 82 percent of participants complete it by the end of April), the assessment of health and social characteristics predated the bulk of participants' spending in a given year. The pattern of results was similar in a sensitivity analysis of spending occurring only in the months after completion of the survey so that all claims-based and survey variables predicted spending in a similarly prospective fashion.

Statistical Analysis

We quantified ACO-level variation in total spending using multilevel linear regression models with ACO random effects to account for sampling error (see Appendix Exhibit A3 for model specification).[12] In all models, we adjusted for demographic characteristics and eligibility categories (age, sex, Medicaid status, disability as original reason for entitlement, and end-stage renal disease), and we adjusted for geography at the Hospital Referral Region (HRR) level (HHR fixed effects) to estimate variation in differences between ACO and non-ACO spending in local service areas. To this base model, we added sets of claims-based and CAHPS survey variables to assess the impact of these further adjustments on variation in ACO spending deviations from local averages. Exhibit 1 summarizes the spending models. Including indicators of the conditions composing the HCC score instead of the summary risk score had a negligible effect on estimates. Models that included health-related variables from the CAHPS survey also included prespecified interactions between these variables. We report standard deviations of differences of ACO spending from non-ACO spending as well as values at the tenth, twenty-fifth, seventy-fifth, and ninetieth percentiles of the distribution of these differences (assuming a normal distribution and following the estimated standard deviation).

Additionally, we examined ACO-level variation in two CAHPS health measures (general health status and mental health status) in similar multilevel models adjusting for demographic characteristics, eligibility categories, and Hospital Referral Region. We then added HCC scores to these models to describe the extent to which current risk adjustment methods account for variation in differences between the general or mental health status of ACO patients and other patients in ACO service areas.

Limitations

Our study had several limitations. First, our sample was necessarily limited to CAHPS survey respondents. The 2011–12 CAHPS samples included 35,532 respondents attributed to ACOs with an average of 142 beneficiaries per ACO (interquartile range, 54–189). Although our statistical approach allowed estimation of ACO variation in spending deviations net of sampling error, using the CAHPS sample limited the precision of our estimates of this variation. Nevertheless, our conclusions were supported by estimates of ACO variation in spending deviations from an analysis of a much larger 20 percent sample of the full Medicare fee-for-service population. Second, nonresponse to the CAHPS survey (response rate of 52.9 percent) led to lower mean spending in the CAHPS sample ($8,683) than in the analogous 20 percent sample of beneficiaries ($9,821), which likely caused us to somewhat underestimate ACO variation in spending deviations. Third, we could only examine the health measures and social factors included in the CAHPS survey when assessing the adequacy of claims-based adjustments. Therefore, we could not discern the extent to which remaining variation in ACO spending deviations reflected variation in efficiency versus variation in other unmeasured patient characteristics, relative to local non-ACO providers.

Study Results

Sample Characteristics

Across all claims-based and survey characteristics, CAHPS participants attributed to ACOs were similar to participants attributed to non-ACO providers, on average, in comparisons adjusted for Hospital Referral Region (Exhibit 2).

Predictors Of Spending

In a model of spending that included demographic characteristics, eligibility categories, HCC score, and CAHPS health measures; statistically significant patient-level predictors included age, disability, presence of end-stage renal disease, HCC score, general health status, mental health status, smoking status, requiring help to complete the survey, and activities of daily living including difficulty with bathing, dressing, eating, walking, getting into and out of chairs, or using the toilet; as well as the interaction between general health status and mental health status (see Appendix Exhibit A4 for full regression results).[12] In summary, each CAHPS health measure except difficulty with one activity of daily living (getting in or out of chairs) significantly predicted spending after adjustment for claims-based measures used by CMS for risk adjustment of ACO benchmarks.

Variation Between ACO Patients And Local Non-ACO Patients

ACOs varied significantly in how their attributed patients' general health status differed from non-ACO patients in ACO service areas. With adjustment for demographic characteristics and eligibility categories, the standard deviation of differences in the percentage of patients in fair or poor health for ACOs versus local non-ACO patients was 2.6 percentage points (Exhibit 3). This standard deviation suggests that the difference in the proportion of patients in fair or poor health between an ACO and the rest of its service area varied from –3.3 percentage points at the tenth percentile of ACOs to 3.3 percentage points at the ninetieth percentile (a 6.6 percentage-point spread relative to an overall mean of 29 percent in fair or poor health). Differences in the proportion of patients reporting their mental health status to be good, fair, or poor varied less markedly--from –1.8 percentage points at the tenth percentile to 1.8 percentage points at the ninetieth percentile after the same adjustments (ACO-level standard deviation, 1.4 percentage points). Further adjustment for HCC score had little effect on this ACO-level variation in general or mental health status, indicating that patients' health status varied across ACOs in ways not accounted for by geography or variables used for risk adjustment in the Medicare ACO programs.

Variation In Differences In Spending

There was substantial variation in differences in total per-beneficiary Medicare spending between ACOs and non-ACO providers in ACO service areas (Exhibit 4). After standard adjustments for demographic characteristics, eligibility categories, and HCC score, ACO spending deviations varied from –$496 (–5.7 percent of mean total spending) at the tenth percentile of ACOs to $456 (5.3 percent) at the ninetieth percentile, as indicated by the standard deviation of $371. Compared with a model adjusting only for demographic characteristics and eligibility categories, the variation in ACO spending deviations was reduced to similar extents by additionally adjusting for HCC score (standard deviation reduced from $453 to $371) or for CAHPS health measures (from $453 to $364). Adding the CAHPS health measures to the model including HCC score reduced the standard deviation further, from $371 to $335, with substantial variation in ACO spending deviations remaining (ranging from –$449 at the tenth percentile of ACOs to $409 at the ninetieth percentile). Combined, HCC score and CAHPS health measures explained 45 percent of variation in ACO spending deviations. Further adjustment for CCW condition indicators and counts reduced the standard deviation by $11, and further adjustment for additional demographic and social characteristics assessed by the CAHPS survey changed the standard deviation minimally (data not shown). The pattern of results was similar in an analysis of eight months of spending following survey completion (Appendix Exhibit A5).[12]

In an analysis of a 20 percent sample of beneficiaries in 2011–12, variation in ACO spending deviations was larger than suggested by our analysis of CAHPS participants ($533 after standard adjustments versus $371 for the CAHPS sample after the same adjustments; data not shown). Mean total per-beneficiary Medicare spending for patients attributed to ACOs was similar to spending for patients attributed to non-ACO providers in ACO service areas ($20 lower for the ACO group; $p = 0.25$; data not shown).

Discussion

In this analysis of 2011–12 CAHPS survey participants and their linked Medicare claims, differences in Medicare spending between ACOs and local non-ACO providers varied greatly. Even after adjustment for patient characteristics assessed from claims and from survey data, ACO spending deviations from local averages exhibited a spread of $858 from the tenth to the ninetieth percentile of ACOs (nearly 10 percent of mean per-beneficiary spending). An analysis of a 20 percent sample of Medicare beneficiaries suggested even greater variation in ACO spending deviations.

Although we could not discern the extent to which remaining variation reflected unmeasured patient factors, these findings suggest wide variation in care efficiency across ACOs, even after accounting for practice patterns and patient demand in ACO service areas. As CMS considers a transition to benchmarks based on regional fee-for-service spending (as in Medicare Advantage),[3] this apparently wide variation in ACO efficiency suggests caution about implementing such measures too rapidly. Even though ACOs with initially high baseline spending may achieve greater initial savings,[4] incentives for ACOs to lower spending are weakened further by spillovers of efforts to lower Medicare spending on fee-for-service revenue from other payers,[5,13] and research to date on the Pioneer model and the Physician Group Practice Demonstration suggests that initial modest savings are unlikely to be compounded year after year,[14–16] although they may eventually grow over time.

Thus, transitioning to a payment model that uses average regional fee-for-service spending as the basis for the benchmark for all ACOs in an area would probably discourage less-efficient organizations from continuing in ACO programs (especially in two-sided risk contracts) if the model were implemented within a few years of participation. Monitoring growth in ACO savings will be critical to determining whether the timing of the transition proposed by CMS will discourage valuable participation by organizations with high baseline spending.

A strategy to promote growth in ACO savings and facilitate a transition to regional benchmarks without discouraging participation would be to eliminate the rebasing of the historical spending component in CMS's proposed benchmarking methodology. That is, the historical benchmark for the initial contract period could be updated for the subsequent contract period to account for regional spending growth and for compositional changes in ACO patients or providers but without rebasing it to reflect the new level of spending achieved by the ACO. Eliminating the rebasing mechanism would allow ACOs more time to recoup the costs of upfront investments in care management infrastructure and sever the link between an ACO's benchmark and its prior savings. In turn this would strengthen incentives for providers to participate and generate savings, thereby fostering more rapid convergence in spending levels between high-spending ACOs and their market averages. Although preliminarily rejected in the proposed regulations, this approach to setting the historical spending component of an ACO's benchmark remains under consideration and could be incorporated into the final rules if supported by analysts and stakeholders.

Once downward pressure on Medicare fee-for-service rates begins under the Merit-Based Incentive Payment System for providers not participating in alternative payment models,[17] significant convergence of initially inefficient ACOs to average local spending may be less important to maintaining their participation under a regional benchmarking model. However, rate reductions under the incentive payment system will not begin until 2019.

In addition, we found that differences in patients' health status between ACOs and local non-ACO providers varied substantially across ACOs after adjustment for HCC scores. Accordingly, additional adjustment for health measures from the CAHPS survey meaningfully reduced the variation in ACO spending deviations. Specifically, the standard deviation was reduced by 10 percent, and the spread between the ninetieth and tenth percentiles was reduced by $94 per beneficiary. These findings suggest that a regional benchmark adjusted only for standard claims-based variables could unfairly penalize ACOs serving sicker patients. Under current ACO rules, ACOs serving sicker patients are not unfairly penalized with inappropriately low benchmarks because an ACO's historical spending serves as the basis for its benchmark.[2] As CMS considers transitioning to a benchmarking methodology that would rely more heavily on risk adjustment to minimize unfair penalties and mitigate incentives for ACOs to select healthier patients, the ACO CAHPS survey could provide valuable additional information for risk adjustment purposes. (The ACO CAHPS survey samples 860 beneficiaries per ACO annually to assess performance on measures of patient experiences and includes the health measures we analyzed from the fee-for-service CAHPS survey as well as additional measures of functional status.[18] Future research could determine how many survey participants per ACO would be required to provide adequate precision for using CAHPS measures in concert with standard claims-based measures for risk adjustment of ACO benchmarks.)

CAHPS survey information on patients also could be used to address incentives for ACOs to upcode to increase risk-adjusted benchmarks. To detect upcoding, survey measures of health status could be compared with concurrent changes in case-mix as assessed by diagnoses in claims. An increase in HCC scores without worsening in self-reported health status, for example, would suggest a change in coding without a true change in the health risks of patients attributed to an ACO. If sufficiently predictive and precise, survey measures could mitigate incentives for upcoding when used in concert with (or eventually perhaps in place of) claims-based measures in risk-adjustment models. Our findings suggest potential for such strategies. Specifically, we found that additional adjustment for CAHPS health measures reduced variation in ACO spending deviations as much as adjustment for HCC scores when compared with estimates from a base model that included only demographic characteristics and eligibility categories.

To the extent that risk adjustment is incomplete or based on lagged data (for example, health status of attributed populations at baseline), an ACO has incentives to serve healthier patients, thereby lowering spending below its benchmark. ACOs can respond to this incentive by improving the health of current patients or attracting different, healthier patients. A fundamental challenge of risk adjustment is that it is difficult to distinguish improvements in health from favorable risk selection (or decrements in health from adverse selection). While better risk adjustment reduces incentives for favorable risk selection, it

may also diminish incentives to improve health; ACOs that improve health status would be penalized with lower benchmarks if benchmarks were adjusted for health status. Thus, it would be important to couple additional adjustment of ACO benchmarks for survey measures of health with stronger incentives for ACOs to perform well on quality measures that assess population health improvement attributable to ACO efforts.

Finally, we found that adjustment for additional social factors assessed by the CAHPS survey and claims-based indicators of conditions from the CCW affected variation in ACO spending deviations minimally. Although we could not ascertain the extent to which residual variation in ACO spending deviations reflected variation in unmeasured patient factors, these findings suggest that much of the remaining variation may reflect variation in efficiency. Thus, to support a transition to regional benchmarks for ACOs, sufficiently robust risk adjustment of ACO benchmarks might be achievable through combined use of claims-based and limited survey information. These findings also suggest that ACOs, with patients attributed through their source of primary care, may pool risk more naturally than hospitals, which have been shown to vary on performance measures to a large extent because of differences in patient characteristics, including many social factors not included in risk adjustment methods used in hospital pay-for-performance programs.[19–21] As ACO programs expand, however, particularly if more safety-net providers participate, it will be important to revisit the sensitivity of ACO benchmarks to adjustment for the social factors we examined.

Conclusion

ACOs vary widely in how their spending levels compare with other providers in ACO service areas. This suggests that measures to equilibrate benchmarks between ACOs with high and low baseline spending--including a transition to regional benchmarks--should be implemented gradually. Monitoring ACO spending deviations and savings will be important to informing optimal implementation of changes to ACO benchmarks. Survey measures of patient health may be useful for limiting ACO gains from upcoding and mitigating the potential unintended consequences of new benchmarking methodologies for ACOs serving sicker patients.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Notes

1. Centers for Medicare and Medicaid Services. Medicare program; Medicare Shared Savings Program: accountable care organizations. Final rule. Fed Regist. 2011; 76(212):67802–990. [PubMed: 22046633]

2. Douven R, McGuire TG, McWilliams JM. Avoiding unintended incentives in ACO payment models. Health Aff (Millwood). 2015; 34(1):143–9. [PubMed: 25561655]

3. Centers for Medicare and Medicaid Services. Medicare program; Medicare Shared Savings Program: accountable care organizations- revised benchmark rebasing methodology, facilitating transition to performance-based risk, and administrative finality of financial calculations; proposed rule. Fed Regist. 2016; 81(22):5824–72.

4. McWilliams JM, Chernew ME, Landon BE, Schwartz AL. Performance differences in year 1 of pioneer accountable care organizations. N Engl J Med. 2015; 372(20):1927–36. [PubMed: 25875195]

5. Chernew, ME., McGuire, TM., McWilliams, JM. Refining the ACO program: issues and options [Internet]. Boston (MA): Harvard Medical School; 2014 Nov 7. [cited 2016 Jan 21]. (Healthcare Markets and Regulations Lab Working Paper). Available from: http://hmrlab.hcp.med.harvard.edu/wp-content/uploads/2014/12/Refining-the-ACO-Program-Issues-and-Options.pdf

6. Geruso, M., Layton, T. Upcoding: evidence from Medicare on squishy risk adjustment [Internet]. Cambridge (MA): National Bureau of Economic Research; 2015 May. [cited 2016 Jan 21]. (NBER Working Paper no. 21222). Available from: http://www.nber.org/papers/w21222

7. Centers for Medicare and Medicaid Services. Fee-for-service (FFS) CAHPS [Internet]. Baltimore (MD): CMS; [last updated 2015 Aug 10; cited 2016 Jan 21]. Available from: https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/CAHPS/ffscahps.html

8. McWilliams JM, Landon BE, Chernew ME, Zaslavsky AM. Changes in patients' experiences in Medicare accountable care organizations. N Engl J Med. 2014; 371(18):1715–24. [PubMed: 25354105]

9. Yun H, Kilgore ML, Curtis JR, Delzell E, Gary LC, Saag KG, et al. Identifying types of nursing facility stays using Medicare claims data: an algorithm and validation. Health Serv Outcomes Res Methodol. 2010; 10(1):100–10.

10. Chronic Conditions Data Warehouse. Welcome to the Chronic Conditions Data Warehouse [home on the Internet]. Baltimore (MD): Centers for Medicare and Medicaid Services; [cited 2016 Jan 21]. Available from: http://www.ccwdata.org

11. Kuye IO, Frank RG, McWilliams JM. Cognition and take-up of subsidized drug benefits by Medicare beneficiaries. JAMA Intern Med. 2013; 173(12):1100–7. [PubMed: 23649604]

12. To access the Appendix, click on the Appendix link in the box to the right of the article online.

13. McWilliams JM, Landom BE, Chernew ME. Changes in health care spending and quality for Medicare beneficiaries associated with a commercial ACO contract. JAMA. 2013; 310(8):829–36. [PubMed: 23982369]

14. Pope G, Kautter J, Leung M, Trisolini M, Adamache W, Smith K. Financial and quality impacts of the Medicare Physician Group Practice Demonstration. Medicare Medicaid Res Rev. 2014; 4(3)

15. Colla CH, Wennberg DE, Meara E, Skinner JS, Gottlieb D, Lewis VA, et al. Spending differences associated with the Medicare Physician Group Practice Demonstration. JAMA. 2012; 308(10):1015–23. [PubMed: 22968890]

16. Nyweide DJ, Lee W, Cuerdon TT, Pham HH, Cox M, Rajkumar R, et al. Association of pioneer accountable care organizations vs. traditional Medicare fee for service with spending, utilization, and patient experience. JAMA. 2015; 313(21):2152–61. [PubMed: 25938875]

17. Rosenthal MB. Physician payment after the SGR--the new meritocracy. N Engl J Med. 2015; 373(13):1187–9. [PubMed: 26398068]

18. Centers for Medicare and Medicaid Services. CAHPS survey for accountable care organizations participating in Medicare initiatives [Internet]. Baltimore (MD): CMS; [cited 2016 Jan 21]. Available from: http://acocahps.cms.gov

19. Barnett ML, Hsu J, McWilliams JM. Patient characteristics and differences in hospital readmission rates. JAMA Intern Med. 2015; 175(11):1803–12. [PubMed: 26368317]

20. Joynt KE, Jha AK. Characteristics of hospitals receiving penalties under the Hospital Readmissions Reduction Program. JAMA. 2013; 309(4):342–3. [PubMed: 23340629]

21. Joynt KE, Orav EJ, Jha AK. Thirty-day readmission rates for Medicare beneficiaries by race and site of care. JAMA. 2011; 305(7):675–81. [PubMed: 21325183]

Biographies

Sherri Rose is an assistant professor of health care policy (biostatistics) in the Department of Health Care Policy at Harvard Medical School, in Boston, Massachusetts.

Alan M. Zaslavsky is a professor of health care policy (statistics) in the Department of Health Care Policy at Harvard Medical School.

J. Michael McWilliams (mcwilliams@hcp.med.harvard.edu) is an associate professor of health care policy and medicine in the Department of Health Care Policy at Harvard Medical School.

EXHIBIT 1

Summary Of Patient Characteristics Included In Different Regression Models

Model description	Demographic and eligibility categories	Geography (hospital referral regions)	HCC score	CAHPS health measures
Demographic	X	X		
Demographic and HCC	X	X	X	
Demographic and CAHPS health measures	X	X		X
Demographic, HCC, and CAHPS health measures	X	X	X	X

SOURCE Authors' analyses of Consumer Assessment of Healthcare Providers and Systems (CAHPS) survey data and linked Medicare claims, 2011–12. NOTES The CAHPS is a survey of fee-for-service Medicare beneficiaries. All models include accountable care organization (ACO) random effects for patients attributed to an ACO (see online Appendix Exhibit A3 for details). (See Note 12 in text.) HCC is Hierarchical Condition Category. CAHPS health measures included general health status, mental health status, requiring help to complete the survey, smoking status, and difficulty with six activities of daily living.

EXHIBIT 2

Hospital Referral Region-Adjusted Comparisons Of Selected Characteristics Between Consumer Assessment Of Healthcare Providers And Systems (CAHPS) Participants Attributed To Accountable Care Organizations (ACOs) And Participants Served By Other Providers

Characteristic	Attributed to an ACO (n = 35,532)	Not attributed to an ACO (n = 160,592)
Medicare claims and enrollment		
Mean age (years)	74	74
Female (%)	55	55
Disabled upon enrollment in Medicare (%)	14	15
Medicaid recipient (%)	5	6
End-stage renal disease (%)	1	1
HCC score	1.3	1.3
Number of CCW conditions	5.9	5.8
CAHPS health measures		
General health status	3.1	3.0
Mental health status	3.8	3.8
Required help to complete survey (%)	10	11
Difficulty with at least one activity of daily living (%)	25	25
CAHPS social measures		
Other insurance (%)	91	90
Among those with other insurance:		
Medigap	19	20
Employer	42	41
Veterans' benefits/Tricare	19	19
Any prescription drug plan	42	42
Race/ethnicity (%)		
Non-Hispanic white	84	84
Education (%)		
Less than high school education	13	14
Lives alone (%)	28	28

SOURCE Authors' analyses of data from the CAHPS survey of fee-for-service Medicare beneficiaries and linked Medicare claims, 2011–12. NOTES Medicaid recipients were not sampled in the 2011 CAHPS survey, leading to a low average proportion with Medicaid. Characteristics were specified in a more detailed fashion in the multilevel regression models (see online Appendix Exhibit A3). (See Note 12 in text.) Activities of daily living included bathing, dressing, eating, walking, using chairs, and using the toilet. HCC is Hierarchical Condition Category. CCW is Chronic Condition Data Warehouse.

EXHIBIT 3

Variation In Differences In Patients' General Health And Mental Health Between Accountable Care Organization (ACO) And Non-ACO Providers in Local Service Areas

	Patient characteristics included in model	Mean percent in poorer health	ACO-level variation in health difference versus local non-ACO providers	Expected distribution of health differences among ACOs			
			Standard deviation (percentage point)	(Percentile)			
				10th	25th	75th	90th
General health status	Demographic	29	2.6	−3.3	−1.8	1.8	3.3
	Demographic and HCC		2.4	−3.1	−1.6	1.6	3.1
Mental health status	Demographic	35	1.4	−1.8	−0.9	0.9	1.8
	Demographic and HCC		1.4	−1.8	−0.9	0.9	1.8

SOURCE Authors' analyses of data from the Consumer Assessment of Healthcare Providers and Systems (CAHPS) survey of fee-for-service Medicare beneficiaries and linked Medicare claims, 2011–12. NOTES For general health status, the poorer health category was "fair" and "poor." For mental health status, this category included "good," "fair," and "poor." Standard deviations were estimated using multilevel regression models. Expected values for percentiles in the ACO distribution of health differences were calculated based on the estimated standard deviations, assuming a normal distribution and using the mean differences in health status for patients attributed to ACOs and those attributed to non-ACO providers (which were close to zero). HCC is Hierarchical Condition Category.

EXHIBIT 4

Variation In Differences In Spending Between Accountable Care Organizations (ACOs) and Non-ACO Providers in Local Service Areas

Patient characteristics included in model	ACO-level variation in spending deviation versus local non-ACO providers	Expected distribution of spending differences among ACOs			
	Standard deviation (in dollars)	(Percentile)			
		10th	25th	75th	90th
Demographic	$453	–$601	–$326	$286	$561
Demographic and HCC	371	–496	–270	230	456
Demographic and CAHPS health measures	364	–487	–266	226	447
Demographic, HCC, and CAHPS health measures	335	–449	–246	206	409

SOURCE Authors' analyses of data from the Consumer Assessment of Healthcare Providers and Systems (CAHPS) survey of fee-for-service Medicare beneficiaries and linked Medicare claims, 2011–12. NOTES Mean total Medicare spending was $8,683 per beneficiary in the CAHPS sample. Standard deviations were estimated using multilevel regression models. Expected values for the percentiles in the ACO distributions were calculated based on the estimated standard deviations, assuming a normal distribution and using the mean difference in total per-beneficiary Medicare spending for patients attributed to ACOs and those attributed to non-ACO providers. HCC is Hierarchical Condition Category.