HSR Health Services Research

© Health Research and Educational Trust
DOI: 10.1111/1475-6773.12829
RESEARCH ARTICLE

Medicare Accountable Care Organizations of Diverse Structures Achieve Comparable Quality and Cost Performance

Leeann N. Comfort, Stephen M. Shortell, Hector P. Rodriguez, and Carrie H. Colla

Objective. To examine whether an empirically derived taxonomy of Accountable Care Organizations (ACOs) is associated with quality and spending performance among patients of ACOs in the Medicare Shared Savings Program (MSSP).

Data Sources. Three waves of the National Survey of ACOs and corresponding publicly available Centers for Medicare & Medicaid Services performance data for NSACO respondents participating in the MSSP ($N = 204$); SK&A Office Based Physicians Database from QuintilesIMS.

Study Design. We compare the performance of three ACO types (physician-led, integrated, and hybrid) for three domains: quality, spending, and likelihood of achieving savings. Sources of performance variation within and between ACO types are compared for each performance measure.

Principal Findings. There is greater heterogeneity within ACO types than between ACO types. There were no consistent differences in quality by ACO type, nor were there differences in likelihood of achieving savings or overall spending per-person-year. There was evidence for higher spending on physician services for physician-led ACOs.

Conclusions. ACOs of diverse structures perform comparably on core MSSP quality and spending measures. CMS should maintain its flexibility and continue to support participation of diverse ACOs. Future research to identify modifiable organizational factors that account for performance variation within ACO types may provide insight as to how best to improve ACO performance based on organizational structure and ownership.

Key Words. Accountable Care Organizations, organizational structure, Medicare Shared Savings Program, risk-based contracts, cost of care

Accountable Care Organizations (ACOs) are a major part of current efforts to improve the health care sector. The ACO model is designed to incentivize high-value care that will reduce spending while improving quality. The model allows groups of providers—physician practices, hospitals, and/or postacute facilities—to join together in assuming financial and medical responsibility for a set of assigned patients. The Medicare Shared Savings Program (MSSP), the largest of the ACO models administered by the Centers for Medicare & Medicaid Services (CMS), encourages organizations to reduce spending and improve quality by developing structures and processes that enable providers to better coordinate care for their fee-for-service patients. Citing the $411 million in savings and improvements in 82 percent of the measures used to assess quality in the ACO program in 2014, the Department of Health & Human Services (HHS) and CMS expanded the ACO model through their sponsorship of the Next Generation ACO and ACO Investment Model programs (HHS 2016).

Even as new payers, specialties, and providers are rapidly adopting the ACO model, little is known about the specific mechanisms that enable ACOs to achieve savings while simultaneously maintaining or improving quality. Pioneer ACOs were found to achieve savings when compared to similar patient populations (McWilliams et al. 2015; Nyweide et al. 2015), as were some MSSP ACOs (Colla et al. 2016; McWilliams et al. 2016). Muhlestein, Saunders, and McClellan (2016) found high levels of quality across MSSP participating ACOs.

However, achieving savings and quality targets has not been uniform across ACOs. In their study of MSSP ACOs, McWilliams et al. (2016) found that ACOs comprised of independent primary care groups achieved larger savings than ACOs with hospital-integrated medical groups. Pioneer ACOs with higher baseline spending levels and those operating in higher-spending geographic areas were also able to achieve greater savings (McWilliams et al. 2015), indicating that ACO programs may most advantage areas and organizations that are in the early stages of their efficiency efforts. Additionally, ACOs that held commercial contracts had higher quality scores, had lower per-person benchmark expenditures, and achieved smaller Medicare savings compared to ACOs that participated only in public Medicare or Medicaid

---

Author for correspondence to Leeann N. Comfort, M.P.P., School of Public Health, University of California, 50 University Hall, Berkeley, CA 94720; e-mail: leeann.comfort@berkeley.edu. Stephen M. Shortell, Ph.D., M.P.H., M.B.A., and Hector P. Rodriguez, Ph.D., M.P.H., are with the School of Public Health, University of California, Berkeley, CA. Carrie H. Colla, Ph.D., is with The Dartmouth Institute for Health Policy and Clinical Practice, Geisel School of Medicine at Dartmouth, Lebanon, NH.

programs (Ouayogodé, Colla, and Lewis 2016; Peiris et al. 2016). A recent analysis by Muhlestein, Saunders, and McClellan (2016) did not find a strong relationship between ACO quality outcomes and net spending, indicating that reductions in spending can be accomplished with little-to-no impact on quality and that organizations that select into ACOs have opportunities to become more efficient.

Some attention has also been given to the financial and quality implications of organizational characteristics at the practice level. Understanding how practice characteristics, including their performance management and improvement activities, affect performance can be helpful for understanding ACO performance since ACOs build upon the existing structures of physician practices. For example, Casalino et al. (2014) found that smaller physician practices have lower rates of potentially preventable hospital admissions and that physician-owned practices had fewer of these admissions than hospital-owned practices. Total per-person spending is also lower in small physician groups compared to medium-sized or large hospital-based groups (McWilliams et al. 2013).

Our research builds on the early effort to provide insights into how ACOs are able to achieve high-quality care at lower cost (e.g., Nyweide et al. 2015; McWilliams et al. 2016; Muhlestein, Saunders, and McClellan 2016; Peiris et al. 2016) by examining the extent to which ACOs of diverse organizational structures, as measured by an empirically derived ACO taxonomy, are systematically associated with performance on outcomes incentivized by the MSSP program. Previous research has provided useful insights by examining individual ACO characteristics in isolation, but focusing on individual characteristics limits the practical use of such research by ACOs. Building upon the ACO taxonomy, our research accounts for the co-occurrence of key ACO characteristics (Shortell et al. 2014; Wu et al. 2016) and examines variability between ACOs according to their MSSP starting year. In the course of our analysis, we also address a measurement challenge not addressed by previous research: Because CMS quality measure specifications vary by ACO starting year and performance year, it is important to compare ACOs using the same measures and measure specifications for a given performance year.

Conceptual Framework

The 923 ACOs in existence as of early 2017 (Muhlestein, Saunders, and McClellan 2017) have diverse organizational structures. Despite the diversity

of these organizations, previous analyses demonstrate high co-occurrence of key ACO characteristics. Shortell et al. (2014) identified three ACO organizational types: (1) the large integrated ACOs, (2) small physician-led ACOs, and the (3) hybrid ACOs that lie between the extremes of the first two. The taxonomy was updated and further validated by Wu et al. (2016). Large integrated ACOs offer the greatest number of services, have fewer percentages of their providers oriented toward primary care, and score lower on physician performance management even though they had greater prior experience in payment reform initiatives. In contrast, the small physician-led ACO cluster tends to offer fewer services, has a greater proportion of primary care providers, and tends to have greater performance management systems in place despite having little prior payment reform experience. The hybrid ACO type reliably falls between the other clusters in many ways. Hybrid ACOs are most often a joint venture by hospitals and physicians; they are smaller than integrated ACOs but larger than physician-led ACOs; they offer more services than the physician-led ACOs but fewer than integrated ACOs. However, hybrid ACOs tend to have proportions of primary care practitioners and physician performance management experience on par with physician-led ACOs.

Using the ACO taxonomy to understand performance differences among ACOs has two advantages: First, it recognizes that there are some commonalities among a diverse set of ACOs. Further, identifying the common clusters of ACO characteristics enables us to analyze the performance variation within and between these clusters in regard to quality and spending. The ability of an ACO to achieve its quality and performance targets, and the strategies ACOs adopt to do so, may be a function of its organizational capabilities and culture. Kreindler et al. (2012) find the ACO model to be amenable to a variety of organizational strategies, which indicates that ACOs may be able to achieve success no matter their organizational type or accompanying improvement strategy. Evidence in physician organizations indicates that performance advantages accrue to larger and more integrated organizations potentially because they use more care management processes designed to lower spending and improve quality—such as use of care managers, electronic disease registries, and sharing quality data with physicians (Rittenhouse et al. 2010; Casalino et al. 2014; Wiley et al. 2015; Rodriguez et al. 2016). These strategies have enabled larger physician organizations to achieve better performance outcomes than their smaller counterparts (Damberg et al. 2010), even while operating at lower per-person costs (Weeks et al. 2010).

In hypothesizing about performance differences between integrated and physician-led ACOs, one must consider the possibility that the ACO types

face different internal and external incentives. For example, ACOs led by hospitals may have less incentive to curb inpatient care if the hospital leadership believes that such losses will not be recouped through shared savings bonuses. Physician-led ACOs, on the other hand, are less likely to be negatively impacted by lower volume of hospital patients and may be better positioned to succeed in the ACO model.

In light of the mixed evidence, our first hypothesis is that large, integrated ACOs may be associated with higher quality but not higher spending when compared to smaller physician-led ACOs. This hypothesis is motivated by the greater use of chronic care management processes and previously identified performance advantages among large physician organizations as well as the possibility that economies of scale enable these organizations to make investments with high fixed costs (like IT infrastructure). However, smaller physician-led organizations have also been found to better prevent hospitalizations for ambulatory care–sensitive conditions (Casalino et al. 2014) and therefore achieve lower per-person spending (McWilliams et al. 2013). Thus, our second hypothesis is that smaller, physician-led ACOs will be more likely to achieve shared savings than larger integrated ACOs. Together, these hypotheses posit that large, integrated ACOs will have higher quality compared to physician-led ACOs, while physician-led ACOs will outperform integrated ACOs on spending measures. Additionally, we presume that the physician-led ACOs will emphasize physician services to achieve quality targets because these services are directly within their control and will therefore have higher per person-year spending for physician services compared to the integrated and hybrid ACO types.

METHODS

Data

We integrated the National Survey of ACOs (NSACO) with the most recent CMS performance and financial data available at the time of this study. The NSACO is an ongoing online survey of ACOs that has been fielded in three waves: October 2012 through May 2013, September 2013 through March 2014, and November 2014 through April 2015. Each ACO is included in only one wave to date; each NSACO wave has been fielded to capture information on new ACO entrants. The response rate across all waves was 64 percent. Previous studies relying on the NSACO data have described the survey and data more fully (e.g., Colla et al. 2014). There was no evidence of nonresponse bias

on key variables when comparing respondent ACOs to the universe of ACOs in the first or second survey administration waves (Colla et al. 2014; Wu et al. 2016). Review of wave three ACOs for the current analysis also showed no statistically significant differences in CMS reported outcomes for survey respondents and nonrespondents. The ACO's annual Herfindahl–Hirschman Index (HHI) market concentrations for specialists and primary care physicians is calculated using the SK&A Office Based Physicians Database provided by IMS Health (now Quintiles).

We report the results of a series of repeated measures analyses that compare the three organizational types over their respective MSSP tenures. Our study sample consists of the 226 NSACO survey respondents that were also included in CMS's August 2015 release of MSSP financial and quality performance data. Because 22 ACOs were dropped due to item nonresponse on key NSACO survey questions, the final analytic sample consisted of 204 MSSP ACOs and 533 performance-year observations. We used unique identifiers to merge the NSACO responses to the organization's corresponding data in the CMS files. This strategy allows us to compare information gathered about the ACO's organizational characteristics shortly after the ACO's establishment with its performance in the subsequent years.

Measures

To assess and categorize each ACO, we used the empirically derived ACO taxonomy developed by Shortell and colleagues (2014) and further validated by Wu et al. (2016). The ACO taxonomy includes eight attributes measured in the NSACO to distinguish each ACOs type: size (number of full-time equivalent (FTE) physicians), breadth of provider participation, scope of services provided, whether the ACO is an integrated delivery system, percent of physician FTEs who are primary care providers, the institutional leadership type, the degree of physician performance management, and payment reform experience (Shortell et al. 2014; Wu et al. 2016). Ouayogodé, Colla, and Lewis (2016) used some of the same variables to examine correlates of ACO performance, but they did not examine differences by ACO types—a more parsimonious way of examining potential associations with performance.

We use the publicly reported quality and financial performance measures released by CMS in 2013, 2014, and 2015. Of the financial data available, we focus on dichotomous indicators for achieving any savings during the performance year, as well as the inpatient, physician services, and total per-person-year spending amounts.

The task of comparing ACOs based on their performance is complicated by yearly changes in the measures used to assess performance. CMS updated its methodology for calculating performance measures each year covered by our analyses. As such, the quality measures and domain items for each performance year will differ slightly according to the year an ACO launched. While these updates are designed to improve the MSSP metrics, they make it difficult to compare ACOs of different starting years or ACOs over time. We address these challenges by measuring each ACO's *relative* performance within each quality domain: patient/caregiver experience, care coordination/patient safety, preventive health, and at-risk population treatment. Relative performance measures were created by identifying each ACO's ventile of performance on a given measure relative to ACOs of the same starting year. Ventiles—each containing 1/20th or 5 percent of the population—were chosen as the most granular level of measurement given the limited number of observations in each combination of starting and performance years. Performance on reverse-coded quality measures was calculated such that lower scores were associated with higher ventiles. Measure-level performance was then averaged across each domain; the measures included in each domain were determined independently for each starting year and performance-year stratum to account for both changes to the measures included in domains and for changes in the methods used to calculate each quality score.

The Herfindahl–Hirschman Index (HHI) was used to account for competition among primary care and specialist providers within an ACO's service area. A detailed account of the methods used to calculate the county-level HHI data is discussed in Fulton (2017). To summarize county-level HHI data at the corresponding state-level service areas publicly reported by CMS, we average the specialist and primary care HHIs for the entire service area, weighted by the number of Medicare fee-for-service beneficiaries in the area (as reported in the corresponding years of the CMS Geographic Variation Public Use File).

Indicators for the geographic region of the United States in which the ACO operates were included the analyses (Northeast, South, Midwest, and West as identified by The Dartmouth Institute's *Dartmouth Atlas*).

Statistical Analyses

While the ACO taxonomy addresses several important operational decisions through proxy measures such as use of performance management

systems and scope of services provided, we also used *t*-tests and chi-square tests to conduct unadjusted comparisons of ACO types for additional measures—such as how the ACO planned to distribute any shared savings bonuses and previous experience with Medicare Advantage and capitated contracts. Pairwise proportions tests, with Bonferroni corrections and Tukey's honest significant difference (HSD) tests designed to identify the specific significant pairwise comparisons within overall statistically significant results, were conducted for significant chi-squared and *F*-test results.

We compared the within-type and between-type variance for each of the performance measures used as dependent variables using a three-level linear restricted maximum-likelihood estimation strategy. The levels for this analysis are performance year observations (level 1) nested within ACOs (level 2) nested within ACO types (level 3), allowing us to identify the extent to which differences in performance are attributable to differences in groups at each level. This analysis adjusts for region, physician HHI, and the year in which the ACO joined the MSSP.

We then leverage the repeated measures of performance year data to estimate a series of a two-level linear restricted maximum-likelihood models, regressing the dependent financial and quality performance variables on the ACO taxonomy types, controlling for region, physician HHI, and the year in which the ACO joined the MSSP. Multilevel models clustered at the individual ACO level enable us to account for unobserved heterogeneity between these organizations—a particularly appealing benefit given the limited number of covariates we can include given the modest ACO sample size.

Sensitivity Analyses

We performed several sensitivity analyses to ensure that the reported results are not artifacts of our model specifications. We assess whether there are systematic differences across years in terms of prior Medicare Advantage or capitated contract experience, or how shared savings bonuses were distributed. As a final robustness check, we compare the results of analyses that pool all ACOs with models for each starting year; while our main specification offers greater statistical power, the year-by-year analysis has the advantage of identifying whether there are statistically significant differences across ACO starting years.

Results

Sample Statistics

Average characteristics are largely indistinguishable between ACO types with only four instances of statistically significant differences between ACO types —and only one difference between integrated and physician-led ACOs (Table 1). We find that a greater proportion of hybrid MSSP ACOs participate in commercial ACOs compared to physician-led ACOs, while a greater proportion of integrated ACOs participate in Medicare Advantage compared to hybrid ACOs. The average HHI for primary care providers is higher for integrated ACOs than for physician-led ACOs (2,157 and 1,723, respectively), indicating that physician-led ACOs tend to operate in primary care markets that are slightly more competitive. Additionally, of the ACOs that report having a plan for distributing shared savings bonuses, integrated ACOs report plans to retain a greater share of the bonus for the ACO body to offset overhead and infrastructure investments in information technology compared to hybrid ACOs (35.6 percent to 19.9 percent, respectively). No other differences between ACO types reached statistical significance.

Sources of Variance

There are no instances when the taxonomy groups account for a substantial portion of the variance on any of the performance measures in the three-level analysis (Table 2): The intraclass correlation coefficient column shows that the maximum proportion variance attributable to the ACO types is 0.05 (for spending on physician services). Instead, the greatest source of performance variation is between ACOs and *within* the ACO types. As such, performance is likely to be a function of organizational characteristics and activities that are present and varying within each type of ACO type.

Quality Outcomes

We do not find evidence to confirm our hypothesis that integrated ACOs demonstrate higher quality than physician-led ACOs when controlling for region, physician HHI, and MSSP starting year (Table 3). In all domains other than patient safety, integrated ACOs achieve higher relative performance compared to physician-led ACOs in adjusted analyses; however, these differences are not statistically significant.

Table 1: Sample Statistics, by ACO Type

	Physician-Led (N = 98)	*Hybrid (N = 38)*	*Integrated (N = 68)*	χ^2	*p*	*PrTest*
Region (No. of ACOs)						
Northeast	22	10	13	(6, 204) = 9.99	.13	—
South	47	12	23			
Midwest	14	12	18			
West	15	4	14			
Previous experience (No. of ACOs)						
Commercial/private ACO	35	23	36	(2, 204) = 8.72	.01	Hybrid > Phys
Medicare Advantage	33	17	37	(2, 122) = 6.17	.05	Int > Hybrid
Capitated contracts	26	12	16	(2, 112) = 1.20	.55	—
	Mean (SD)	*Mean (SD)*	*Mean (SD)*	*F*	*p*	η^2	*HSD*
Beneficiary characteristics (Percent of ACO population)							
Disabled	14.5 (7.2)	15.5 (6.9)	16.4 (9.8)	(2, 206) = 1.18	.31	0.01	—
End-stage renal disease	1.1 (0.7)	1.0 (0.7)	1.0 (0.6)	(2, 206) = 0.40	.67	0.00	—
Medicaid eligible	8.6 (11.2)	6.2 (6.4)	7.9 (10.1)	(2, 206) = 0.79	.46	0.01	—
Female	53.1 (4.8)	54.8 (4.6)	54.7 (4.8)	(2, 193) = 2.90	.06	0.03	—
Race/ethnicity							
Black	9.4 (12.1)	7.9 (7.4)	7.9 (7.9)	(2, 193) = 0.55	.58	0.01	—
Hispanic	2.6 (5.4)	2.1 (3.8)	1.9 (3.4)	(2, 193) = 0.49	.61	0.01	—
Asian	1.6 (6.6)	0.9 (1.3)	2.4 (8.8)	(2, 193) = 0.57	.57	0.01	—
Native American	0.2 (0.5)	0.1 (0.3)	0.2 (0.5)	(2, 193) = 0.52	.59	0.01	—
Other	1.9 (1.4)	1.9 (1.0)	2.4 (2.0)	(2, 193) = 1.96	.14	0.02	—

Continued

Table 1 *Continued*

	Mean (SD)	*Mean (SD)*	*Mean (SD)*	*F*	*p*	η^2	*HSD*
HHI (Average HHI score)							
Average All Providers	2,426 (1,019)	2,580 (1,080)	2,885 (1,136)	(2, 205) = 3.69	.03	0.04	—
Primary care	1,723 (746)	1,987 (926)	2,157 (932)	(2, 205) = 5.48	.01	0.05	Int > Phys
Specialty care	3,130 (1,406)	3,173 (1,379)	3,614 (1,469)	(2, 205) = 2.48	.09	0.02	—
ACOs with plan for distributing bonus (percent of ACOs)	75.0	68.8	75.5				
Distribution of bonus							
Kept by ACO administrative body	27.3 (18.9)	19.9 (17.2)	35.6 (27.8)	(2, 102) = 3.47	.04	0.07	Int > Hybrid
Shared with member organizations	28.6 (27.6)	43.2 (33.8)	30.6 (24.8)	(2, 83) = 1.75	.18	0.04	—
Shared directly with physicians	44.6 (30.0)	34.7 (26.7)	37.9 (28.9)	(2, 95) = 0.95	.39	0.02	—
Shared with external partners	7.4 (12.5)	7.5 (15.0)	12.2 (16.8)	(2, 74) = 0.90	.41	0.02	—
Other	4.4 (19.0)	4.8 (12.4)	7.5 (18.6)	(2, 59) = 0.19	.82	0.01	—

Notes. PrTest: Pairwise test of proportions with Bonferroni corrections to identify pairwise significant differences; η^2: the proportion of variance attributable to ACO types; Tukey's honest significant difference (HSD) test: identifies significant pairwise comparisons within overall statistically significant results.

Table 2: Sources of Variance

	Variance between ACO Types	*ICC*	*Variance between ACOs*	*ICC*	*Residual Variance*	*N*
Quality (Ventile of performance)						
Patient experience	0.27	0.01	2.37***	.554	2.14***	491
Patient safety	6.16×10^{-10}	4.14×10^{-20}	2.43***	.642	1.81***	493
Preventive health	6.49×10^{-9}	3.18×10^{-18}	2.98***	.673	2.08***	484
Diabetes	2.28×10^{-10}	3.31×10^{-21}	3.09***	.607	2.49***	484
Hypertension	4.99×10^{-5}	8.61×10^{-11}	4.17***	.601	3.40***	484
Ischemic vascular disease	0.46	0.01	2.65***	.324	3.88***	484
Heart failure	3.92×10^{-10}	4.62×10^{-21}	3.02***	.275	4.91***	480
Coronary artery disease	1.29×10^{-8}	6.69×10^{-18}	3.14***	.399	3.86***	484
Financial outcomes						
Overall spending ($)	332.4	0.01	2,070.7***	.440	2,365.0***	533
Physician services ($)	167.8	0.05	742.2***	.964	146.6***	533
Inpatient services ($)	4.8×10^{-5}	2.50×10^{-15}	935***	.939	238.8***	533
Any savings achieved (0/1)	2.4×10^{-7}	7.13×10^{-15}	2.19***	.593	—	493

Notes. ICC: Intraclass correlation coefficient, proportion of total variance shared among units in same cluster (Taxonomy and ACO).
***$p<.001$.

Financial Outcomes

We find no evidence to refute our hypothesis that large, integrated ACOs may be associated with higher quality but not higher spending when compared to smaller physician-led ACOs. As Table 4 shows, there are no significant differences between ACO types in total per person-year spending or inpatient spending per person-year. In fact, the significant differences in physician spending indicate that integrated ACOs have *lower* spending in this category compared to physician-led ACOs. Higher physician services spending among physician-led ACOs is consistent with our assumption that physicians are likely to increase use of ambulatory care services in the pursuit of lower overall spending through reduced inpatient utilization for ambulatory care–sensitive conditions.

Despite differences in spending by ACO type, we find no differences in the likelihood of achieving any savings. As the final column of Table 4 shows, the odds of achieving any savings are no greater for integrated ACOs compared to physician-led ACOs.

Table 3: Quality Results by Domain: Performance Ventile

	Patient Experience	*Patient Safety*	*Preventive Services*	*Diabetes*	*Hypertension*	*Ischemic Vascular Disease*	*Heart Failure*	*Coronary Artery Disease*
Taxonomy (Reference: "physician-led")								
Integrated	0.68 (0.47)	−0.26 (0.47)	0.07 (0.57)	0.41 (0.61)	0.61 (0.81)	1.20 (0.63)	0.39 (0.76)	0.72 (0.70)
Hybrid	0.77 (0.57)	−0.26 (0.56)	−0.04 (0.68)	0.22 (0.72)	1.35 (0.97)	0.52 (0.75)	−0.80 (0.90)	0.49 (0.83)
Region (Reference: "northeast")								
South	−1.52** (0.53)	−0.13 (0.52)	−1.27* (0.63)	−1.45* (0.67)	−2.87** (0.91)	−1.16 (0.70)	1.57 (0.84)	0.14 (0.77)
Midwest	−0.74 (0.64)	−0.24 (0.62)	−0.86 (0.76)	−0.37 (0.81)	−0.00 (1.09)	0.32 (0.85)	0.42 (1.02)	0.69 (0.93)
West	−3.69*** (0.67)	1.64* (0.66)	−3.05*** (0.81)	−3.35*** (0.86)	−4.99*** (1.16)	−3.85*** (0.90)	−0.01 (1.08)	−2.09* (0.99)
Starting year (Reference: 2012)								
2013	−0.16 (0.48)	0.07 (0.47)	−0.14 (0.58)	−0.12 (0.61)	−0.79 (0.82)	0.19 (0.63)	0.48 (0.75)	0.05 (0.69)
2014	0.21 (0.51)	0.11 (0.50)	−0.48 (0.61)	−0.22 (0.66)	−0.73 (0.88)	0.03 (0.69)	−0.49 (0.83)	−0.33 (0.76)
Competition								
Physician HHI	0.00 (0.00)	0.00** (0.00)	0.00 (0.00)	0.00*** (0.00)	−0.00 (0.00)	0.00* (0.00)	−0.00 (0.00)	−0.00 (0.00)
Constant	10.91*** (0.57)	9.25*** (0.55)	11.00*** (0.67)	10.08*** (0.72)	12.55*** (0.98)	9.66*** (0.79)	10.21*** (0.95)	10.86*** (0.86)
Random part								
SD(ACO)	2.37***	2.44***	3.00***	3.11***	4.17***	2.65***	3.02***	3.15***
SD(Residual)	2.14***	1.81***	2.08***	2.49***	3.40***	3.88***	4.91***	3.86***
Observations	491	493	484	484	484	484	480	484

Notes. SEs in parentheses; performance ventiles calculated using ACOs of same starting and performance year; *p<.05, **p<.01, ***p<.001.

Table 4: Spending Results, by ACO Type

	Total Spending (per person-year)	*Physician Services (per person-year)*	*Inpatient Services (per person-year)*	*Achieved Savings (odds ratio)*
Taxonomy (Reference: "physician-led")				
Integrated	−829.71 (426.04)	−339.40** (123.64)	−155.23 (156.92)	1.21 (0.60)
Hybrid	−547.50 (498.95)	−296.86* (146.59)	−112.28 (185.94)	1.13 (0.67)
Region (Reference: "Northeast")				
South	−656.04 (481.51)	−28.20 (139.98)	−352.76* (177.67)	3.91* (2.20)
Midwest	−837.98 (570.38)	−463.46** (160.93)	−252.67 (204.65)	1.03 (0.68)
West	230.76 (590.41)	277.89 (172.54)	179.69 (218.90)	0.52 (0.37)
Starting year (Reference: 2012)				
2013	−1,685.91*** (419.46)	−139.16 (124.68)	102.43 (158.09)	0.89 (0.44)
2014	−1,607.64*** (470.14)	40.98 (134.50)	90.70 (170.75)	0.85 (0.46)
Competition				
Physician HHI	−0.31* (0.16)	−0.14*** (0.01)	−0.05* (0.02)	1.00 (0.00)
Constant	13,667.40*** (534.88)	3,883.22*** (134.59)	3,593.59*** (172.66)	0.93 (0.57)
Random part				
SD(ACO)	2,076.93***	742.57***	937.15***	2.19***
SD(Residual)	2,363.36***	146.59***	238.81***	—
Observations	533	533	533	493

Notes. Reference group: physician-led ACOs; all models include region and Herfindahl–Hirschman Index (HHI) market concentrations for physicians; models with all years combined also include indicators for ACO starting year.
*p<.05, **p<.01, ***p<.001.

Sensitivity Analyses

The sensitivity analyses ruled out the possibility that the results were driven by several potential omitted variables: There were no systematic differences across ACO types and years in terms of prior Medicare Advantage or capitated contract experience, or how shared savings bonuses were distributed among ACO participants (Table 1). Likewise, we found that patient populations—in terms of demographics, concurrent Medicaid eligibility, and Medicare eligibility based on disability status or End-Stage Renal Disease diagnosis—did not systematically differ between ACO types. Our results did not differ when we added these variables to the regression model.

Our comparison of the analyses that pool all ACOs with models for each starting year shows no statistically significant differences across ACO starting years. There is some variability between ACOs of different starting years, but these differences are not statistically significant (see Figure S1, which depicts the coefficient plots from the robustness checks that compare our original model that combines all years with separate models for each starting year).

Discussion

Previous research identified three major types or clusters of ACOs based on eight conceptually and empirically derived attributes (Shortell et al. 2014; Wu et al. 2016). Performance on quality or financial measures in the first 3 years of the MSSP was not associated with ACO type. In fact, ACO type accounts for a maximum proportion of 5 percent of the variance in performance (Table 2, ICC columns). As such, each ACO type appears equally capable of achieving quality performance, which is consistent with early ACO research by Kreindler et al. (2012), who found that the strategies employed to achieve integration varied according to the unique characteristics and needs of participants. This suggests that CMS and other payers should continue to support all types of ACOs while focusing attention on improving the effectiveness of the lowest performers in order to reduce performance variation within ACO types.

In an early step toward disentangling the heterogeneity in ACO performance, Peiris et al. (2016) found that Medicare ACOs that also held commercial ACO contracts tended to achieve smaller savings but had higher quality scores than those ACOs that did not participate in commercial payer ACO contracts. Such findings are consistent with other research that found noticeable differences between ACOs that contracted with public and private payers, such as greater experience with pay-for-performance initiatives and greater reliance on upfront care management payments for those ACOs contracting with private payers (Lewis et al. 2014). However, our sensitivity analyses found that our conclusions are largely unchanged by the inclusion of commercial contract information. Even though we find that hybrid ACOs are more likely to participate in commercial ACO contracts compared to physician-led ACOs, this trend does not lead to advantages (or disadvantages) in the quality or cost performance of ACOs.

ACO types are neither predictive of performance nor proxies for differences in production decisions, highlighting the importance of better understanding the sources of heterogeneity in performance within ACO types. Sources of within-ACO-type performance variation may include institutional pressures (Goodrick and Reay 2016) faced in particular by the larger integrated group model ACOs; high transaction costs (Mick and Shay 2016) faced by the smaller physician-led ACOs; the need for each type to implement high reliability care management processes (Vogus and Singer 2016) and to accelerate learning of what works and does not work (Nembhard and Tucker 2016).

Insights on how to drive improvement within each ACO type can be gained by focusing on how professionals understand and apply the concept of integration within their organization. For example, Kreindler et al. (2012) found that ACO leaders understood integration in very different ways: some perceived it as more of an administrative formality while others strongly identified as partners with their new ACO colleagues. Providers' commitment to the activities of coordinated care may be a function of how they conceptualize integration and the responsibilities it entails; participation in care management tasks is likely driven more by personal motivation and buy-in than by top-down reform efforts. Notions of what integration entails differ by ACO type, so future research examining the ways in which ACOs integrate services based on their organizational structure may aid in identifying the drivers of within-ACO type variance in quality and cost performance. Conceptualizations of integration among ACO stakeholders may also be related to their willingness to undertake performance-enhancing initiatives such as implementing new patient communication technologies and dedicating resources to programs for patient education and engagement (Shortell et al. 2017). Clarifying the ways leaders design and execute their organizational integration strategies —and the extent to which providers buy-in to the approaches—can provide important insight into the differences in performance across organizations that are structurally similar.

LIMITATIONS

There are several limitations to our study, spanning both internal and external validity concerns. While our study uses repeated observations for the performance data that serve as dependent variables, the full NSACO survey data were collected from each site only once. There is clear need for longitudinal analyses of survey-based information as ACOs evolve, and we look forward to leveraging future follow-up surveys to continue this work. We were also limited in the number of independent variables that could be examined due to a modest sample size. Third, while our results may be useful for analyzing Medicare ACOs, the lessons learned here might not be directly applicable to ACOs operating outside of the Medicare Shared Savings Program. Commercial ACOs are more numerous and cover more lives than federal ACOs on average (Muhlestein et al. 2016), so additional research is needed to assess whether the patterns we identified are exhibited among commercial ACOs.

Conclusions

Patterns in quality and financial performance are generally not consistent across diverse ACO types. In analyzing the association of ACO type and performance on key quality and financial metrics, we find a substantial amount of performance heterogeneity within ACO types. As such, we recommend future research on ACO performance focus on previously unexamined factors that may account for variance within types, such as potential differences in organizational culture, leadership, use of team-based care, and different strategies to achieve organizational integration and how such differences interact with ACO structure and resources to influence performance. Interorganizational learning between ACOs of the same type could enhance performance of individual ACOs by creating improvement strategies suited to the unique needs of the organization. Such collaborative learning exercise may also be of interest to policy makers who seek to improve overall ACO performance.

Our analyses also indicate that researchers and policy makers should be cautious when pooling ACOs of different implementation dates and performance years due to changes in measure specification, especially in high-stakes performance-based financial incentive programs. While our robustness checks demonstrated no statistically significant differences in our analyses by starting year (Figure S1), the variations in coefficient plots underscore the importance of continued monitoring of associations between ACO type and performance. Finally, given we find a great deal of heterogeneity within ACO types, future research should identify the factors affecting organizational performance within as well as potentially between ACOs types.

Acknowledgments

Joint Acknowledgment/Disclosure Statement: The authors thank Benjamin Usadi of The Dartmouth Institute and Daniel Arnold of the University of California, Berkeley, for their integral contributions to the data preparation and statistical analysis for this study. Earlier versions of this analysis were presented at the 2016 AcademyHealth Annual Research Meeting, the 2017 Berkeley-Stanford Organizational Behavior Conference, and the 2016 National Research Service Award Trainee's Conference. The authors thank the participants at this conference for their helpful feedback.

This study was supported by the Agency for Healthcare Research and Quality (AHRQ) Comparative Health System Performance Initiative under Grant #1U19HS024075, which studies how health care delivery systems promote evidence-based practices and patient-centered outcomes research in delivering care, as well as the National Institutes of Health (Grant #R33AG044251) and The Commonwealth Fund (Grant #20160616). The National Survey of Accountable Care Organizations was supported by The Commonwealth Fund (Grant #20130084) and AHRQ (Grant #1T32HS022241-01). Ms. Comfort received support from the AHRQ under the Ruth L. Kirschstein National Research Service Award Institutional Research Training Grant (#4T32HS022241). The contents of this study are solely the responsibility of the authors and do not necessarily represent the official views of AHRQ.

Disclosures: None.

Disclaimer: At the time of the manuscript's preparation, Dr. Shortell was a member of the Advisory Board for the Centene Corporation—Medicaid Managed Care Company (St. Louis, MO).

References

Casalino, L. P., M. F. Pesko, A. M. Ryan, J. L. Mendelsohn, K. R. Copeland, P. P. Ramsay, X. Sun, D. R. Rittenhouse, and S. M. Shortell. 2014. "Small Primary Care Physician Practices Have Low Rates of Preventable Hospital Admissions." *Health Affairs* 33 (9): 1680–8. https://doi.org/10.1377/hlthaff.2014.0434.

Colla, C. H., V. A. Lewis, S. M. Shortell, and E. S. Fisher. 2014. "First National Survey of ACOs Finds that Physicians Are Playing Strong Leadership and Ownership Roles." *Health Affairs* 33 (6): 964–71. https://doi.org/10.1377/hlthaff.2013.1463.

Colla, C. H., V. A. Lewis, L.-S. Kao, A. J. O'Malley, C.-H. Chang, and E. S. Fisher. 2016. "Association between Medicare Accountable Care Organization Implementation and Spending among Clinically Vulnerable Beneficiaries." *JAMA Internal Medicine* 176 (8): 1167–75. https://doi.org/10.1001/jamainternmed.2016.2827.

Damberg, C. L., S. M. Shortell, K. Raube, R. R. Gillies, D. Rittenhouse, R. K. McCurdy, L. P. Casalino, and J. Adams. 2010. "Relationship between Quality Improvement Processes and Clinical Performance." *American Journal of Managed Care* 16: 601–6. Available at http://www.ajmc.com/journals/issue/2010/2010-08-vol16-n08/AJMC_10aug_Damberg_601to606/

Department of Health & Human Services (HHS). 2016. "New Hospitals and Health Care Providers Join Successful, Cutting-Edge Federal Initiative That Cuts Costs and Puts Patients at the Center of Their Care," Press Release HHS.gov [accessed

on October 13, 2016]. Available at http://www.hhs.gov/about/news/2016/01/11/new-hospitals-and-health-care-providers-join-successful-cutting-edge-federal-initiative.html.

Fulton, B. D. 2017. "Health Care Market Concentration Trends in The United States: Evidence and Policy Responses." *Health Affairs* 36 (9): 1530–8. https://doi.org/10.1377/hlthaff.2017.0556.

Goodrick, E., and T. Reay. 2016. "An Institutional Perspective on Accountable Care Organizations." *Medical Care Research and Review* 73 (6): 685–93. https://doi.org/10.1177/1077558716641832.

Kreindler, S. A., B. K. Larson, F. M. Wu, K. L. Carluzzo, J. N. Gbemudu, A. Struthers, A. D. V. Citters, S. M. Shortell, E. C. Nelson, and E. S. Fisher. 2012. "Interpretations of Integration in Early Accountable Care Organizations." *The Milbank Quarterly* 90 (3): 457–83.

Lewis, V. A., C. H. Colla, W. L. Schpero, S. M. Shortell, and E. S. Fisher. 2014. "ACO Contracting with Private and Public Payers: A Baseline Comparative Analysis." *The American Journal of Managed Care* 20 (12): 1008–14.

McWilliams, J. M., M. E. Chernew, A. M. Zaslavsky, P. Hamed, and B. E. Landon. 2013. "Delivery System Integration and Health Care Spending and Quality for Medicare Beneficiaries." *JAMA Internal Medicine* 173 (15): 1447. https://doi.org/10.1001/jamainternmed.2013.6886.

McWilliams, J. M., M. E. Chernew, B. E. Landon, and A. L. Schwartz. 2015. "Performance Differences in Year 1 of Pioneer Accountable Care Organizations." *New England Journal of Medicine* 372 (20): 1927–36. https://doi.org/10.1056/NEJMsa1414929.

McWilliams, J. M., L. A. Hatfield, M. E. Chernew, B. E. Landon, and A. L. Schwartz. 2016. "Early Performance of Accountable Care Organizations in Medicare." *New England Journal of Medicine* 374 (24): 2357–66. https://doi.org/10.1056/NEJMsa1600142.

Mick, S. S. F., and P. D. Shay. 2016. "Accountable Care Organizations and Transaction Cost Economics." *Medical Care Research and Review* 73 (6): 649–59. https://doi.org/10.1177/1077558716640411.

Muhlestein, D., R. Saunders, and M. McClellan. 2016. "Medicare Accountable Care Organization Results For 2015: The Journey to Better Quality and Lower Costs Continues". Health Affairs Blog. Available at http://healthaffairs.org/blog/2016/09/09/medicare-accountable-care-organization-results-for-2015-the-journey-to-better-quality-and-lower-costs-continues/.

———. 2017. Growth of ACOs and Alternative Payment Models In 2017. Health Affairs Blog. Available at http://healthaffairs.org/blog/2017/06/28/growth-of-acos-and-alternative-payment-models-in-2017/.

Muhlestein, D., T. Tu, K. de Lisle, and T. Merrill. 2016. "Hospital Participation in ACOs Associated with Other Value-Based Program Improvement." *The American Journal of Managed Care* 22 (7): e241–8.

Nembhard, I. M., and A. L. Tucker. 2016. "Applying Organizational Learning Research to Accountable Care Organizations." *Medical Care Research and Review* 73 (6): 673–84. https://doi.org/10.1177/1077558716640415.

Nyweide, D. J., W. Lee, T. T. Cuerdon, H. H. Pham, M. Cox, R. Rajkumar, and P. H. Conway. 2015. "Association of Pioneer Accountable Care Organizations vs Traditional Medicare Fee for Service with Spending, Utilization, and Patient Experience." *Journal of the American Medical Association* 313 (21): 2152. https://doi.org/10.1001/jama.2015.4930.

Ouayogodé, M. H., C. H. Colla, and V. A. Lewis. 2016. "Determinants of Success in Shared Savings Programs: An Analysis of ACO and Market Characteristics." *Healthcare* 5: 53–61. Available at http://linkinghub.elsevier.com/retrieve/pii/S2213076416300793, https://doi.org/10.1016/j.hjdsi.2016.08.002.

Peiris, D., M. C. Phipps-Taylor, C. A. Stachowski, L.-S. Kao, S. M. Shortell, V. A. Lewis, M. B. Rosenthal, and C. H. Colla. 2016. "ACOs Holding Commercial Contracts are Larger and More Efficient Than Noncommercial ACOs." *Health Affairs* 35 (10): 1849–56. https://doi.org/10.1377/hlthaff.2016.0387.

Rittenhouse, D. R., S. M. Shortell, R. R. Gillies, L. P. Casalino, J. C. Robinson, R. K. McCurdy, and J. Siddique. 2010. "Improving Chronic Illness Care: Findings from a National Study of Care Management Processes in Large Physician Practices." *Medical Care Research and Review* 67 (3): 301–20. https://doi.org/10.1177/1077558709353324.

Rodriguez, H. P., S. R. McClellan, S. Bibi, L. P. Casalino, P. P. Ramsay, and S. M. Shortell. 2016. "Increased Use of Care Management Processes and Expanded Health Information Technology Functions by Practice Ownership and Medicaid Revenue." *Medical Care Research and Review* 73 (3): 308–28. https://doi.org/10.1177/1077558715613233.

Shortell, S. M., F. M. Wu, V. A. Lewis, C. H. Colla, and E. S. Fisher. 2014. "A Taxonomy of Accountable Care Organizations for Policy and Practice." *Health Services Research* 49 (6): 1883–99. https://doi.org/10.1111/1475-6773.12234.

Shortell, S. M., B. Y. Poon, P. P. Ramsay, H. P. Rodriguez, S. L. Ivey, T. Huber, J. Rich, and T. Summerfelt. 2017. "A Multilevel Analysis of Patient Engagement and Patient-Reported Outcomes in Primary Care Practices of Accountable Care Organizations." *Journal of General Internal Medicine* 32 (6): 640–7. https://doi.org/10.1007/s11606-016-3980-z.

Vogus, T. J., and S. J. Singer. 2016. "Creating Highly Reliable Accountable Care Organizations." *Medical Care Research and Review* 73 (6): 660–72. https://doi.org/10.1177/1077558716640413.

Weeks, W. B., D. J. Gottlieb, D. J. Nyweide, J. M. Sutherland, J. Bynum, L. P. Casalino, R. R. Gillies, S. M. Shortell, and E. S. Fisher. 2010. "Higher Health Care Quality and Bigger Savings Found at Large Multispecialty Medical Groups." *Health Affairs* 29 (5): 991–7. https://doi.org/10.1377/hlthaff.2009.0388.

Wiley, J. A., D. R. Rittenhouse, S. M. Shortell, L. P. Casalino, P. P. Ramsay, S. Bibi, A. M. Ryan, K. R. Copeland, and J. A. Alexander. 2015. "Managing Chronic Illness: Physician Practices Increased the Use of Care Management and Medical Home Processes." *Health Affairs* 34 (1): 78–86. https://doi.org/10.1377/hlthaff.2014.0404.

Wu, F. M., S. M. Shortell, V. A. Lewis, C. H. Colla, and E. S. Fisher. 2016. "Assessing Differences between Early and Later Adopters of Accountable Care Organizations Using Taxonomic Analysis." *Health Services Research* 51: 2318–29. https://doi.org/10.1111/1475-6773.12473.

Supporting Information

Additional supporting information may be found online in the supporting information tab for this article:

Appendix SA1: Author Matrix.

Figure S1: Coefficient Plots, All Years Combined and by ACO Starting Year.