Annals of Internal Medicine ORIGINAL RESEARCH

Performance in the Medicare Shared Savings Program After Accounting for Nonrandom Exit

An Instrumental Variable Analysis

Adam A. Markovitz, BS; John M. Hollingsworth, MD, MS; John Z. Ayanian, MD, MPP; Edward C. Norton, PhD; Phyllis L. Yan, MS; and Andrew M. Ryan, PhD

Background: Accountable care organizations (ACOs) in the Medicare Shared Savings Program (MSSP) are associated with modest savings. However, prior research may overstate this effect if high-cost clinicians exit ACOs.

Objective: To evaluate the effect of the MSSP on spending and quality while accounting for clinicians' nonrandom exit.

Design: Similar to prior MSSP analyses, this study compared MSSP ACO participants versus control beneficiaries using adjusted longitudinal models that accounted for secular trends, market factors, and beneficiary characteristics. To further account for selection effects, the share of nearby clinicians in the MSSP was used as an instrumental variable. Hip fracture served as a falsification outcome. The authors also tested for compositional changes among MSSP participants.

Setting: Fee-for-service Medicare, 2008 through 2014.

Patients: A 20% sample (97 204 192 beneficiary-quarters).

Measurements: Total spending, 4 quality indicators, and hospitalization for hip fracture.

Results: In adjusted longitudinal models, the MSSP was associated with spending reductions (change, −$118 [95% CI, −$151 to −$85] per beneficiary-quarter) and improvements in all 4 quality indicators. In instrumental variable models, the MSSP was not associated with spending (change, $5 [CI, −$51 to $62] per beneficiary-quarter) or quality. In falsification tests, the MSSP was associated with hip fracture in the adjusted model (−0.24 hospitalizations for hip fracture [CI, −0.32 to −0.16 hospitalizations] per 1000 beneficiary-quarters) but not in the instrumental variable model (0.05 hospitalizations [CI, −0.10 to 0.20 hospitalizations] per 1000 beneficiary-quarters). Compositional changes were driven by high-cost clinicians exiting ACOs: High-cost clinicians (99th percentile) had a 30.4% chance of exiting the MSSP, compared with a 13.8% chance among median-cost clinicians (50th percentile).

Limitation: The study used an observational design and administrative data.

Conclusion: After adjustment for clinicians' nonrandom exit, the MSSP was not associated with improvements in spending or quality. Selection effects—including exit of high-cost clinicians—may drive estimates of savings in the MSSP.

Primary Funding Source: Horowitz Foundation for Social Policy, Agency for Healthcare Research and Quality, and National Institute on Aging.

Ann Intern Med. 2019;171:27-36. doi:10.7326/M18-2539 **Annals.org**
For author affiliations, see end of text.
This article was published at Annals.org on 18 June 2019.

Accountable care organizations (ACOs) are arguably the most widespread and far-reaching value-based reform in the United States, with more than 900 ACO contracts covering more than 32 million lives (1). A leading example is the Medicare Shared Savings Program (MSSP), in which groups of clinicians, hospitals, and other providers voluntarily assume responsibility for the spending and quality outcomes of a defined population of fee-for-service Medicare beneficiaries. Evidence suggests that MSSP ACOs are associated with modest improvements in spending and quality (2–4).

However, MSSP evaluations may be subject to confounding from nonrandom participation or attrition within ACOs. Clinicians and provider groups may be recruited to or choose to join ACOs because of their desire or ability to deliver high-quality, efficient care (5). These providers may simultaneously engage in other payment reforms, such as Medicare Advantage (MA), that could lower spending (6, 7). Accountable care organizations may also selectively drop high-cost clinicians and their patient panels to reduce measured spending and earn shared savings (8, 9). Despite these selective pressures operating at the clinician, provider group, and MSSP contracting levels, research to date has not identified the degree to which selection bias may affect evaluations of the MSSP.

In this study, we used national Medicare data to evaluate changes in spending and quality performance while accounting for selection effects in the MSSP. As in previous analyses, we estimated adjusted longitudinal models that account for secular trends, market-level factors, and observed differences across MSSP participants and local control beneficiaries. To account for clinicians' nonrandom entry to and exit from the MSSP, we used the share of clinicians participating in the MSSP within a 50-mile radius of a given clinician's practice location as an instrumental variable. We also tested for changes in the composition of MSSP participants over time and whether these changes drive estimates of spending in the program.

See also:

Web-Only
Supplement

METHODS

Data Sources and Study Population

We analyzed national claims data from 2008 through 2014 for a random 20% sample of Medicare fee-for-service beneficiaries aged 65 years or older. Participation in an MSSP was defined using the ACO Provider- and Beneficiary-level files from the Centers for Medicare & Medicaid Services (CMS); these files list the beneficiaries, clinicians, and provider groups participating in the MSSP. Following the MSSP specifications given by CMS (10), we attributed ACO beneficiaries to the provider group and clinician (defined by Taxpayer Identification Number) within their ACO from whom they received the plurality of eligible evaluation and management services. To improve comparability to prior work (2, 3), we restricted analyses to beneficiaries who could be attributed using claims submitted by a primary care clinician in the outpatient setting (**Supplement**, available at Annals.org).

As specified by the MSSP (10), in each year we excluded beneficiaries who were ineligible for attribution to an MSSP ACO, were enrolled in MA, or were not continuously enrolled in Medicare Parts A and B. We also excluded beneficiaries attributed to ACOs with no listed clinician participants in the MSSP Provider-file or who resided outside a hospital referral region.

Study Variables

Primary and Secondary Outcomes

Our primary study outcome was total price-standardized Medicare spending per beneficiary per quarter. Price standardization accounted for variation resulting from regional wage indices and payments for indirect medical education, disproportionate share hospitals, and new technologies (12).

We included the following secondary outcomes: component spending for inpatient, outpatient, professional, and skilled-nursing facility services; clinical quality performance, including indicators for diabetes (glycated hemoglobin testing, low-density lipoprotein cholesterol testing, and diabetic retinal examinations) and mammography; and hospital use, including rates of all-cause hospitalization, preventable hospitalization (for any 1 of 11 ambulatory care–sensitive conditions) (13), all-cause 30-day readmissions (14), and emergency department visits.

As a falsification test, we evaluated the effect of the MSSP on hospitalization for hip fracture (15). We selected hip fracture because it is a marker of health status that is unlikely to be affected by medical practice or ACOs in the short term. Also, because hip fracture requires hospitalization, it is less subject to changes in diagnostic intensity or treatment preferences that may be correlated with ACO participation (16). Previous research has shown that rates of hip fracture hospitalization are a sensitive indicator of true population incidence (17, 18).

Exposure

We defined beneficiary attribution to the MSSP using a time-varying indicator that equaled 1 if the beneficiary was attributed to an MSSP ACO in a given quarter and otherwise equaled 0. This captured the staggered entry of ACOs into MSSP contracts (April 2012, July 2012, January 2013, and January 2014) and the ability of participants to enter and exit ACOs each year. In light of reports that MSSP savings are concentrated among early entrants (2–4), we also evaluated spending changes across ACOs' year of entry into MSSP contracts.

Instrumental Variable: MSSP Supply

We hypothesized that clinicians practicing in closer proximity to other MSSP providers would be more likely to join and remain in MSSP ACOs but would otherwise be similar to local clinicians. Our instrument is conceptually similar to the commonly used "differential distance" instrument (19–21).

In each quarter, MSSP supply was defined as the ratio of MSSP clinicians to all clinicians within a 50-mile radius of a given clinician's practice location, after discounting the supply of clinicians who were further away within the 50-mile radius (22) (**Supplement Figure 1**, available at Annals.org). The discount factor was estimated by modeling how the likelihood of 2 clinicians participating in the same MSSP ACO varied by the inverse distance between those 2 clinicians (22, 23) and was allowed to vary across urban, large rural, small rural, and isolated ZIP codes (Methods section of the **Supplement** and **Supplement Table 1**, available at Annals.org) (24–26). We tested the robustness of the instrumental variable using MSSP supply restricted to a radius of 10, 25, 50, 75, or 150 miles, with or without discounting. In contrast to standard longitudinal models, instrumental variable models capture treatment effects among marginal clinicians—that is, those participating in the MSSP because of greater exposure to MSSP supply.

Statistical Analysis

Performance in the MSSP

We estimated 2 sets of linear regression models for each spending and quality outcome. First, we estimated adjusted longitudinal models comparing concurrent changes in spending and quality between MSSP participants and local control beneficiaries within the same health care market, as defined by hospital referral region (2–4, 11, 27, 28). Models included a quarterly indicator for MSSP attribution, market fixed effects, year fixed effects, seasonal indicators, and beneficiary characteristics. We adjusted for beneficiary age, sex, race/ethnicity, disability, end-stage renal disease, dual eligibility for Medicaid, and CMS Hierarchical Condition Category risk score, as well as area-level poverty and education (29).

We then estimated instrumental variable models that were analogous to the adjusted longitudinal model but used MSSP supply as an instrumental variable. By integrating MSSP supply into a longitudinal framework, these models specifically used within-market changes in MSSP supply to identify the effect of the MSSP, thus accounting for secular trends and

fixed differences across markets that could confound the relationship between MSSP supply and local performance changes. The Methods section of the Supplement provides model specification details.

As a falsification test, we estimated the effect of the MSSP on the rate of hospitalization for hip fracture using the adjusted longitudinal model and the instrumental variable model. Because the MSSP should not affect hip fracture rates during our study period, observed associations would suggest the presence of residual confounding not addressed by the given statistical model.

Compositional Changes

To identify the degree to which changes in the composition of MSSP participants drive performance estimates, we estimated changes in spending and hip fracture using supplemental longitudinal models that included market-year fixed effects, beneficiary fixed effects, market-year and ACO fixed effects (with ACOs defined by groups of Taxpayer Identification Numbers ultimately forming MSSP ACOs) (2, 3, 30), or clinician fixed effects (Methods section of the Supplement). The model including market-year and ACO fixed effects is most similar to the models used in the principal MSSP evaluations (2, 3). We tested instrumental variable robustness in a similar manner, estimating changes in spending and hip fracture in models that included market-year fixed effects; market-year and ACO fixed effects; or MSSP supply restricted to a radius of 10, 25, 50, 75, or 150 miles, with or without discounting.

In supplemental analyses, we evaluated 2 potential mechanisms underlying changes in the composition of MSSP participants or their practice environment. First, we assessed whether MSSP ACOs strategically prune providers according to spending performance in the program. Specifically, we tested whether average spending by clinicians or provider groups in the MSSP predicted their exit from an MSSP ACO in the subsequent year. We also tested whether a beneficiary whose primary clinician exited the MSSP was more likely to exit than a beneficiary whose clinician remained in the program. To distinguish nonrandom attrition from random churn of participants in and out of the MSSP, we examined the association between clinician or provider group spending and entry into the MSSP. For these analyses, we used models that adjusted for average beneficiary characteristics, year fixed effects, and market fixed effects. We also performed a series of intention-to-treat analyses to determine the extent to which nonrandom attrition of clinicians or beneficiaries biased estimates of the MSSP's effect (Methods section of the Supplement).

Second, we analyzed whether MA penetration (a proxy for managed care) was associated with MSSP penetration at the county level and whether adjustment for county-level MA penetration affected savings estimates of ACOs. If such adjustment reduced savings estimates, it would suggest that other time-varying changes occurring alongside ACOs may bias standard estimates.

All analyses specified robust SEs to account for clustering at the market level. Statistical analyses were performed using Stata, version 15.1 (StataCorp). The University of Michigan Institutional Review Board exempted our study from review.

Role of the Funding Source

This research is supported by grants from the Horowitz Foundation for Social Policy, the Agency for Healthcare Research and Quality, and the National Institute on Aging. The funders had no role in the analysis presented here.

RESULTS

MSSP Participation and Spending

Medicare fee-for-service beneficiaries contributed 97 204 192 beneficiary-quarters from 2008 through 2014 (6 871 934 unique beneficiaries). Participants in the MSSP included 835 100 beneficiaries, 30 331 clinicians, 6883 provider groups, and 337 ACOs from 2012 through 2014. Average Medicare spending per beneficiary-quarter was $2335 (SD, $7312).

Instrument Validity

We began by confirming the validity of MSSP supply as an instrumental variable. First, MSSP supply was highly correlated with MSSP participation ($F_{1,306}$ statistic = 444, where instruments with F statistics above 10 are considered strong) (31, 32). Second, beneficiary covariates were extremely well balanced across MSSP supply (Table). Third, spending trends in the pre-MSSP period were similar for beneficiaries with high versus low MSSP supply (differential trend, −$2 [95% CI, −$4 to −$1] per beneficiary-quarter) (Supplement Figure 2 and Supplement Table 2, available at Annals.org).

In contrast, across observed MSSP status, MSSP beneficiaries were less likely to be dually eligible or disabled, have substantial comorbidity, or live in areas with high poverty or low educational attainment (Table). Spending trends in the pre-MSSP period differed substantially between beneficiaries who did and did not eventually enter the MSSP (differential trend, $17 [CI, $16 to $19] per beneficiary-quarter) (Supplement Figure 3 and Supplement Table 2, available at Annals.org), violating the parallel trends assumption. These findings suggest that unobserved sources of selection bias were more evenly distributed across MSSP supply than across observed MSSP status.

Changes in Spending

Figure 1 shows the association between the MSSP and spending. In the adjusted longitudinal model, the MSSP was associated with a reduction in total spending (change, −$118 [CI, −$151 to −$85] per beneficiary-quarter). Savings in the adjusted model were due to reductions in inpatient services (change, −$58 [CI, −$73 to −$44] per beneficiary-quarter), outpatient services (change, −$15 [CI, −$24 to −$7] per beneficiary-quarter), and skilled-

Table. Characteristics of Beneficiaries and Clinicians Across MSSP ACO Participation and MSSP Supply*

Characteristic	Covariate Balance Across MSSP ACO Status, Pre-ACO Period†			Covariate Balance Across MSSP Supply (Instrumental Variable), Pre-ACO Period‡		
	Unadjusted Mean Value for Control Beneficiaries (*n* = 47 948 364 Beneficiary-Quarters)†	Unadjusted Mean Value for MSSP Beneficiaries (*n* = 8 368 351 Beneficiary-Quarters)†	Standardized Difference	Unadjusted Mean Value for Low MSSP Supply (*n* = 32 627 168 Beneficiary-Quarters)‡	Unadjusted Mean Value for High MSSP Supply (*n* = 23 689 544 Beneficiary-Quarters)‡	Standardized Difference
Patient characteristics						
Age, *y*	76.8	75.4	−0.20	76.6	76.6	0.00
Female sex, %	60.2	60.9	0.01	60.3	60.4	0.00
Race/ethnicity, %§						
Non-Hispanic white	85.9	87.0	0.03	85.4	86.9	0.05
Non-Hispanic black	6.9	6.3	−0.03	7.0	6.7	−0.01
Hispanic	4.1	3.6	−0.03	4.3	3.7	−0.03
Other	3.1	3.2	0.00	3.4	2.7	−0.04
Dual eligibility for Medicaid, *mo*‖	1.5	1.1	−0.12	1.4	1.4	−0.01
HCC risk score¶	1.1	1.0	−0.22	1.1	1.1	0.01
Disability, %**	8.8	7.2	−0.06	8.5	8.6	0.01
End-stage renal disease, %**	0.7	0.3	−0.06	0.7	0.7	0.00
Area-level characteristics, %						
Below federal poverty level††	13.9	12.3	−0.19	13.8	13.6	−0.03
With high school degree††	86.4	87.8	0.17	86.5	86.8	0.04
With college degree††	28.0	31.3	0.20	28.6	28.4	−0.02
Enrolled in Medicare Advantage‡‡	21.9	21.5	−0.04	21.9	21.9	0.00
Patient outcomes						
Total spending per beneficiary-quarter, $	2510	17 253	−0.12	2379	2413	0.00
Hip fractures per 1000 beneficiary-quarters, *n*§§	2.3	1.2	−0.03	2.2	2.2	0.00

ACO = accountable care organization; HCC = Hierarchical Condition Category; MSSP = Medicare Shared Savings Program.
* Descriptive characteristics were based on the pre-MSSP period (January 2008–March 2012) and were calculated for all beneficiaries appearing in ≥1 quarter of the post-MSSP period, because this was required for determining whether the beneficiary's primary clinician was exposed to high vs. low MSSP supply (*n* = 56 316 714 beneficiary-quarters).
† For descriptive analyses, MSSP participants were defined as beneficiaries who ever participated in the MSSP over the study period; control beneficiaries were defined as beneficiaries who never participated in the MSSP over the study period. For regression analyses of the effect of the MSSP on study outcomes, MSSP participation was defined using a time-varying indicator that equaled 1 if the beneficiary was attributed to the MSSP in a given quarter and otherwise equaled 0.
‡ High vs. low MSSP supply was defined by whether the attributed clinician's median MSSP value was above (1 = high) or equal to or below (0 = low) the median MSSP supply value for the hospital referral region over the entire post-MSSP period. The beginning of the post-MSSP period was designated as the first quarter in which the hospital referral region contained ≥1 clinician participating in the MSSP.
§ Determined from Medicare enrollment files and based on the Research Triangle Institute race/ethnicity designation.
‖ Defined by the number of months in the year that the beneficiary's state of residence paid the monthly premium for Part B coverage (range, 0–12 mo).
¶ Calculated using Medicare demographic and diagnostic data from the prior year's enrollment and claims files. Higher scores indicate higher predicted spending in the present year.
** Defined by original reason for Medicare entitlement, regardless of current status.
†† These proportions were defined at the ZIP code tabulation area using American Community Survey data.
‡‡ This proportion was defined at the county level using enrollment data for a 20% random sample of Medicare beneficiaries. Beneficiaries were defined as participating in Medicare Advantage if they were enrolled for ≥1 mo during the year and were otherwise defined as not enrolled in Medicare Advantage.
§§ Defined as an acute care hospital claim with a primary diagnosis of hip fracture (820.xx) in a given quarter.

nursing facility services (change, −$41 [CI, −$55 to −$27] per beneficiary-quarter). In the instrumental variable model, the MSSP was not associated with changes in total spending (change, $5 [CI, −$51 to $62] per beneficiary-quarter). The instrumental variable estimate for spending differed significantly from the adjusted estimate (difference-in-Sargan test $P < 0.001$) (**Supplement Table 3**, available at Annals.org) (33). Estimated savings were smaller in instrumental variable models than in adjusted models across each ACO cohort (**Supplement Figure 4**, available at Annals.org).

Falsification Test

In the adjusted longitudinal model, the MSSP was associated with a substantial decrease in the falsification outcome, hip fracture (change, −0.24 hospitalizations [CI, −0.32 to −0.16 hospitalizations] per 1000 beneficiary-quarters) (**Figure 2**). This represents an 11% decrease in the rate of hip fracture hospitalization (0.24 fewer hospitalizations per 1000 beneficiary-quarters divided by an adjusted mean rate of 2.12 hospitalizations per 1000 beneficiary-quarters). In contrast, the MSSP was not associated with hip fracture hospitalizations in

the instrumental variable model (change, 0.05 hospitalizations [CI, −0.10 to 0.20 hospitalizations] per 1000 beneficiary-quarters). The instrumental variable estimate for hip fracture differed significantly from the adjusted estimate (difference-in-Sargan test $P = 0.002$) (Supplement Table 3) (33).

Effects of Compositional Change

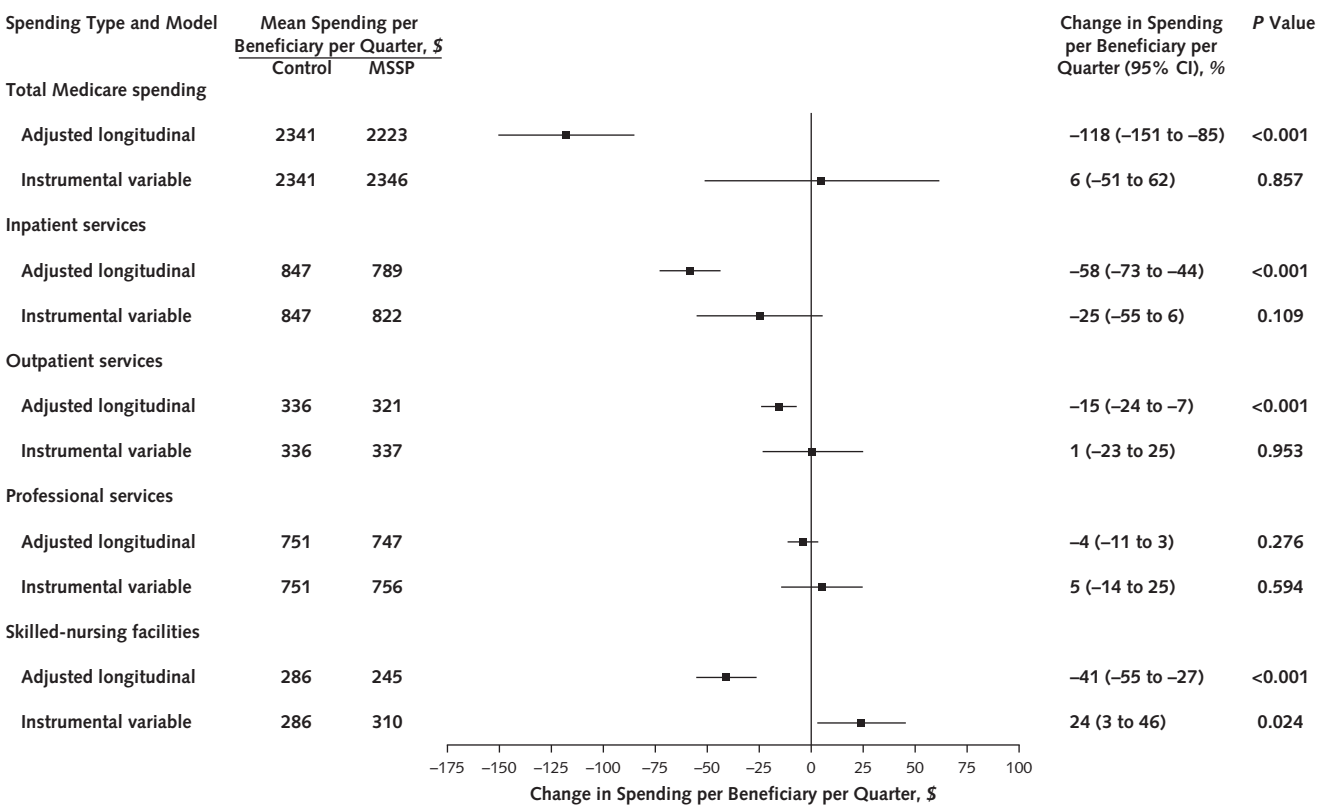
Figure 2 shows the large influence of compositional changes on estimates of the MSSP's effect. The association between the MSSP and hip fracture (Figure 2, *top*) consistently decreased in models that controlled for fixed differences across market years (change, −0.28 hospitalizations [CI, −0.37 to −0.20 hospitalizations] per 1000 beneficiary-quarters), beneficiaries (change, −0.20 hospitalizations [CI, −0.28 to −0.12 hospitalizations] per 1000 beneficiary-quarters), market years and ACOs (change, −0.17 hospitalizations [CI, −0.24 to −0.10 hospitalizations] per 1000 beneficiary-quarters), and clinicians (change, −0.06 hospitalizations [CI, −0.11 to −0.01 hospitalizations] per 1000 beneficiary-quarters). The association between the MSSP and spending (Figure 2, *bottom*) followed a similar pattern: Estimates of total savings decreased in models that accounted for fixed differences across market years (−$125 per beneficiary-quarter), beneficiaries (−$90 per beneficiary-quarter), market years and ACOs (−$66 per beneficiary-quarter), and clinicians (−$28 per beneficiary-quarter) (all $P < 0.001$).

In contrast, the MSSP was not associated with spending or hip fracture in the instrumental variable model (Figure 2). This finding was robust across a wide range of specifications, including instrumental variable models using market-year fixed effects; market-year and ACO fixed effects; or MSSP supply restricted to a radius of 25, 50, 75, or 150 miles, with or without discounting (Supplement Figure 5 and Supplement Table 4, available at Annals.org). Although changes in MSSP and MA penetration were modestly associated at the county level (Supplement Figure 6, available at Annals.org), adjustment for MA penetration did not affect ACO estimates (Supplement Table 4).

Changes in Quality and Hospital Use

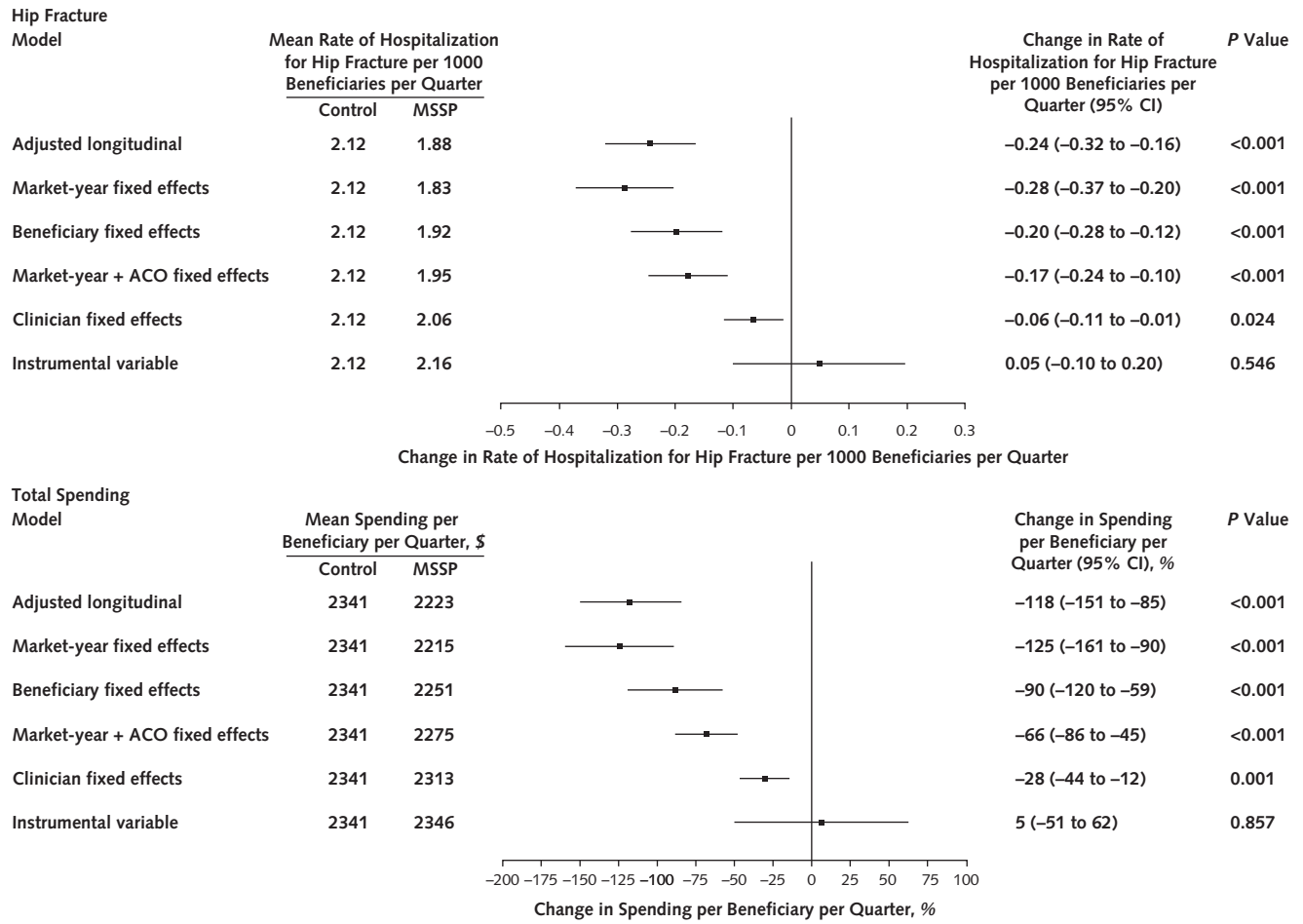
The MSSP was associated with improvements in all 4 clinical quality indicators in the adjusted longitudinal model but not in the instrumental variable model (Figure 3 and Supplement Table 3). The MSSP was associated with modest decreases in all-cause hospitalizations and preventable hospitalizations in the adjusted longitudinal model but not in the instrumental variable model; conversely, reductions in all-cause 30-day read-

Figure 1. Changes in Medicare spending for beneficiaries attributed to MSSP ACOs versus control beneficiaries.



Mean outcomes for control beneficiaries were derived from the adjusted longitudinal model. Total spending was the sum of Medicare spending for inpatient, outpatient, professional, and skilled-nursing facility services. Component spending was defined by claims for services from the following research identifiable files: Medicare Provider Analysis and Review (for inpatient and skilled-nursing facility services); carrier (for professional services); and outpatient services (for outpatient services). ACO = accountable care organization; MSSP = Medicare Shared Savings Program.

Figure 2. Changes in hospitalization and spending for hip fracture across models for fixed differences across MSSP participants versus control beneficiaries.



Mean outcomes for control beneficiaries were derived from the adjusted longitudinal model. Hospitalization for hip fracture was defined as a Medicare acute-care hospital claim for a primary diagnosis of hip fracture (820.xx). Total spending was the sum of Medicare spending for inpatient, outpatient, professional, and skilled-nursing facility services. ACO = accountable care organization; MSSP = Medicare Shared Savings Program. **Top.** Change in rate of hospitalization for hip fracture per 1000 beneficiary-quarters. **Bottom.** Change in total spending per beneficiary-quarter.

missions and emergency department visits were observed in both models (**Supplement Figure 7** and **Supplement Table 3**, available at Annals.org).

Clinician Attrition From MSSP ACOs

Between 2013 and 2014, a total of 4054 clinicians in our sample exited the MSSP. **Figure 4** shows the relationship between a clinician's average spending in the MSSP and subsequent exit from the program. Clinicians in the 95th and 99th percentiles of average spending had a 22.3% and 30.4% chance, respectively, of exiting the MSSP in the next year, compared with a 13.8% chance among clinicians with median (50th percentile) spending (risk difference, 8.5 and 16.6 percentage points, respectively; both $P < 0.001$) (**Supplement Table 5**, available at Annals.org). Entry into ACOs was also modestly higher among clinicians in the 99th percentile of prior spending than among median-spending clinicians (risk difference, 3.8 percentage points; $P < 0.001$) (**Figure 4** and **Supplement Table 5**). Conversely, MSSP entry and exit by provider groups was less common and not associated with groups' prior spending (**Supplement Figure 8**, available at Annals.org).

The chance of exiting the MSSP was 21.6 percentage points higher for beneficiaries whose primary clinician exited the MSSP than for those whose clinicians remained in the MSSP (31.1% vs. 9.5%; risk difference $P < 0.001$) (**Supplement Figure 9** and **Supplement Table 6**, available at Annals.org). Even among beneficiaries whose clinicians remained in the MSSP, however, the highest-cost beneficiaries (99th percentile) had a greater chance of exiting the MSSP than median-spending beneficiaries (17.7% vs. 8.9%; risk difference, 8.8 percentage points; $P < 0.001$) (**Supplement Figure 9** and **Supplement Table 6**).

In supplemental intention-to-treat analyses that directly addressed attrition of clinicians and beneficiaries from the MSSP, the MSSP was not associated with

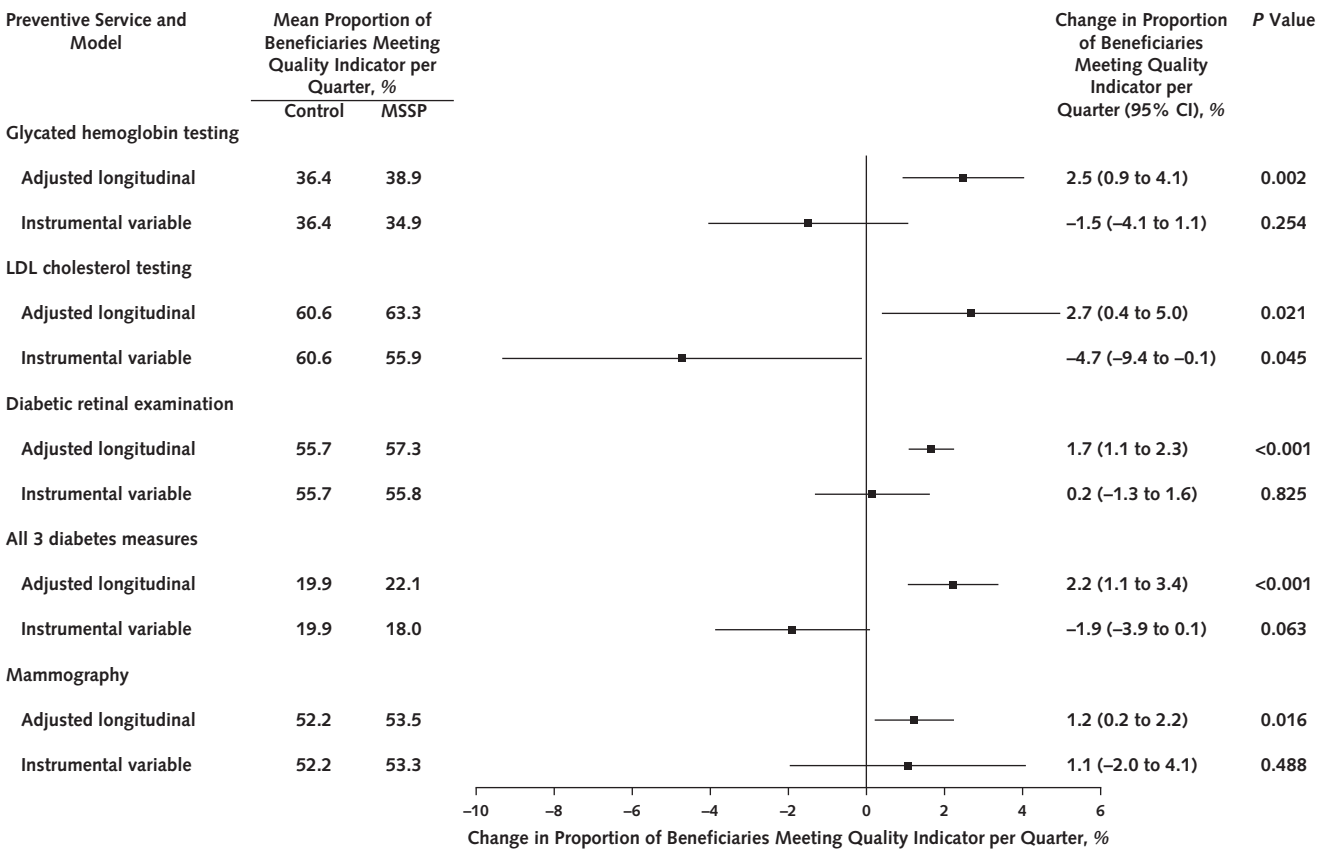
changes in spending or hip fracture (Supplement Table 4). This provides further evidence that nonrandom attrition of high-cost clinicians and high-cost beneficiaries drives estimated savings in the MSSP.

DISCUSSION

Participation in MSSP ACOs was not associated with improvements in spending, quality, or hip fracture (the falsification outcome) in our instrumental variable analysis. Adjusted longitudinal models—similar to those used by other studies to estimate the effect of ACOs—failed our falsification test, suggesting residual confounding from unobserved severity of patient illness. Before the start of the MSSP, spending trends also differed between beneficiaries who did and did not eventually enter the program. Supplemental analyses found that high-cost clinicians and beneficiaries were disproportionately likely to exit MSSP ACOs. Together, our results suggest that improved quality and spending performance in this voluntary program may have been driven by nonrandom exit of clinicians and their patient panels from the MSSP.

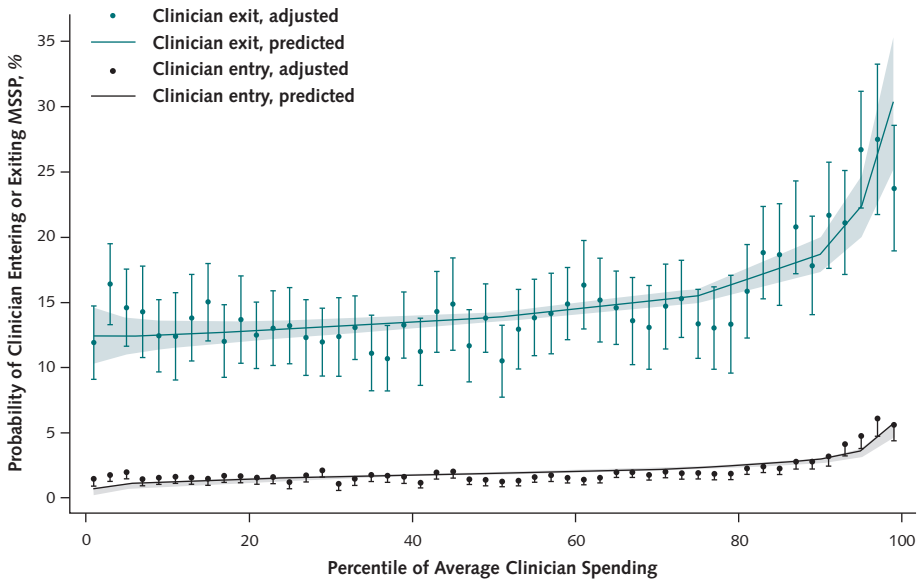
Our conclusion that the MSSP was not associated with improvements in spending, quality, or most measures of hospital use differs from that of previous evaluations of Medicare ACOs (2–4, 11, 27, 28). Our instrumental variable model addresses selection effects not directly captured in previous evaluations. One might consider addressing selection bias in the MSSP with standard difference-in-differences models that account for fixed differences across ACOs. However, the influence of compositional changes within MSSP ACOs is highlighted by the consistent reduction in savings and hip fracture estimates that occurred with progressively greater adjustment for fixed differences across MSSP participants and control beneficiaries. Instrumental variable models may account for compositional changes by evaluating spending and quality among clinicians who join and remain in the MSSP because of their proximity to other MSSP clinicians rather than their record of spending performance in the program.

Figure 3. Changes in clinical quality performance for beneficiaries attributed to MSSP ACOs versus control beneficiaries.



The proportions of MSSP beneficiaries and control beneficiaries receiving the preventive service are given in percent. Mean outcomes for control beneficiaries were derived from the adjusted longitudinal model. Diabetes clinical indicators were derived from National Quality Forum specifications and included glycated hemoglobin testing (in quarter of interest), LDL cholesterol testing (in quarter of interest or previous 3 quarters), and diabetic retinal examination (in quarter of interest or previous 3 quarters). Analyses were limited to beneficiaries with diabetes (n = 15 323 604 beneficiary-quarters). The mammography indicator was derived from National Quality Forum specifications. Beneficiaries were defined as meeting the indicator if they received a mammogram (in quarter of interest or previous 7 quarters). Mammography analyses were limited to female beneficiaries aged 65–69 y (n = 11 922 514 beneficiary-quarters). We excluded all observations following a mastectomy. ACO = accountable care organization; LDL = low-density lipoprotein; MSSP = Medicare Shared Savings Program.

Figure 4. Association between clinician spending and probability of clinician exiting or entering the MSSP.



Average clinician spending was defined as the average Medicare spending per beneficiary per year of the clinician's attributed patient panel in the 3 y before MSSP exit or entry determination. The probability of a clinician entering or exiting the MSSP was estimated as a function of the clinician's average spending, average beneficiary characteristics of the clinician's attributed patient panel, market fixed effects, and year fixed effects. Quadratic and cubic spending terms were included to allow for any potential nonlinearities in the effect of spending performance on MSSP participation. Analyses of MSSP exit (n = 21 418 clinician-years) were restricted to clinicians participating in the MSSP during the year before analysis. Analyses of MSSP entry (n = 161 957 clinician-years) were restricted to clinicians not participating in the MSSP during the year before analysis and did not include ACO formation, i.e., participation in an ACO's first contract year. Both sets of analyses were restricted to ACOs that entered MSSP contracts in 2012 or 2013, because 2012–2014 MSSP data could not be used to determine clinician exit or entry for ACOs formed in 2014. ACO = accountable care organization; MSSP = Medicare Shared Savings Program.

Our finding that high-cost clinicians and their patient panels disproportionately exit the MSSP is consistent with prior studies showing that clinicians and beneficiaries with higher predicted spending are more likely to leave Medicare ACOs (8, 9). We found no evidence that MSSP ACOs recruit low-cost clinicians. This may be because ACOs can more readily observe the spending of clinicians and beneficiaries already in their ACO than that of those who have not yet joined.

Pruning high-cost clinicians from ACO contracts could have large effects on spending estimates and may contribute to reported findings that MSSP savings grow over time (3). For example, dropping clinicians in the top 95th or 99th percentiles of spending from the MSSP would lower average ACO spending per beneficiary in our sample by 3.1% or 0.6%, respectively (**Supplement Table 7**, available at Annals.org). On the other hand, average ACO spending per beneficiary in our sample would increase by 1.1% if we included clinicians who exited the MSSP and excluded those who entered (**Supplement Table 8**, available at Annals.org). These changes are similar in magnitude to previous estimates of total savings in the MSSP (2–4, 27, 28). Our supplemental intention-to-treat analyses provide further evidence that nonrandom attrition of high-cost clinicians and high-cost beneficiaries drives estimated savings in the MSSP (**Supplement Table 4**).

Possible limitations of our analysis merit discussion. First, some unobserved factor may confound the relationship between MSSP supply and performance. Markets with greater MSSP supply likely contain more clinicians capable of implementing value-based initiatives. However, because our instrumental variable models used only within-market changes in MSSP supply over time, the instrumental variable is not subject to this concern. Moreover, in contrast to the adjusted longitudinal model, MSSP supply had excellent covariate balance and parallel spending trends and was not associated with changes in the falsification outcome. Second, instrumental variable models typically generate less precise estimates than adjusted models, risking type II error. However, instrumental variable estimates of spending, quality, and hip fracture differed significantly from adjusted longitudinal estimates (33). Third, changes in hip fracture hospitalization, our falsification outcome, may represent true improvements to underlying beneficiary health by ACOs. However, this is unlikely during our brief post-MSSP period and in light of the relatively modest health improvements reported in other ACO analyses (2–4, 11, 27, 28). Fourth, because we attributed beneficiaries to MSSP ACOs using CMS's official lists of MSSP beneficiaries and clinicians, our results may differ from initial evaluations that used their own attribution methods (2–4, 30). However, assessing whether MSSP ACOs strategically recruit or drop high-cost clinicians required that we use CMS's official lists. Fifth, as with prior MSSP evaluations, our study may be affected by growth in MA participation over time. However, controlling for MA penetration did not alter our estimates of spending. Sixth, health care

use outside ACOs by attributed beneficiaries may attenuate the effect of ACOs. Among beneficiaries assigned to an ACO, 15.6% of primary care visits and 71.9% of specialist visits are with non-ACO clinicians (**Supplement Table 9**, available at Annals.org). However, these patterns of care are common to all ACO evaluations and should have no bearing on our findings that savings are driven by selective attrition of high-cost clinicians from ACOs.

Finally, because instrumental variable models estimate treatment effects among marginal clinicians, our findings may not generalize to clinicians who would either always join the MSSP (for example, because they are employed by an early adopter health system) or never join the MSSP (for example, because they lack the desire or capacity to undertake such a reform). At the same time, this population—clinicians joining the MSSP because of greater engagement with and access to other MSSP clinicians—is of intrinsic policy relevance. That the MSSP had little effect among these participants is pertinent for assessing the potential economic and health rewards of governmental policies intended to promote ACO participation, such as the Medicare Access and CHIP Reauthorization Act of 2015 (34).

Our results challenge the view that MSSP ACOs have lowered spending and improved quality; they indicate that savings by MSSP ACOs may be driven by nonrandom exit of high-cost clinicians and their patient panels from this voluntary program. These findings suggest caution in extending ACOs to other settings and patients without stronger evidence that the program saves money or improves quality of care. Our study underscores the degree to which selection bias may affect evaluations of voluntary reforms and the challenges inherent in evaluating these programs.

From University of Michigan Medical School and School of Public Health, Ann Arbor, Michigan (A.A.M.); University of Michigan Medical School and Institute for Healthcare Policy and Innovation, Ann Arbor, Michigan (J.M.H.); University of Michigan Medical School, School of Public Health, Gerald R. Ford School of Public Policy, and Institute for Healthcare Policy and Innovation, Ann Arbor, Michigan (J.Z.A.); University of Michigan School of Public Health and Institute for Healthcare Policy and Innovation, Ann Arbor, Michigan, and National Bureau of Economic Research, Cambridge, Massachusetts (E.C.N.); University of Michigan Medical School, Ann Arbor, Michigan (P.L.Y.); and University of Michigan School of Public Health, Center for Evaluating Health Reform, Institute for Healthcare Policy and Innovation, Ann Arbor, Michigan (A.M.R.).

Note: Dr. Ryan had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Financial Support: Mr. Markovitz is supported by the Horowitz Foundation for Social Policy and by grant R36HS025615 from the Agency for Healthcare Research and Quality. Dr. Hollingsworth is supported by grants R01HS024728 and 1R01HS024525-01A1 from the Agency for Healthcare Research and Quality. Dr. Ryan is supported by grant R01AG047932 from the National Institute on Aging.

Disclosures: Dr. Hollingsworth reports other support from University of California, San Francisco, that had no effect on the analysis presented in this manuscript. Authors not named here have disclosed no conflicts of interest. Disclosures can also be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M18-2539.

Reproducible Research Statement: *Study protocol:* Available in the **Supplement** (available at Annals.org). *Statistical code:* Available at https://github.com/adammarkovitz. *Data set:* Available for purchase from ResDAC.

Corresponding Author: Andrew M. Ryan, PhD, Department of Health Management and Policy, University of Michigan School of Public Health, M3124 SPH II, 1415 Washington Heights, Ann Arbor, MI 48109-2029; e-mail, amryan@umich.edu.

Current author addresses and author contributions are available at Annals.org.

References

1. **Muhlestein D, Saunders RS, McClellan MB.** Growth of ACOs and alternative payment models in 2017. Health Affairs Blog. 28 June 2017. Accessed at www.healthaffairs.org/do/10.1377/hblog20170628.060719/full on 1 April 2019.
2. **McWilliams JM, Hatfield LA, Chernew ME, Landon BE, Schwartz AL.** Early performance of accountable care organizations in Medicare. N Engl J Med. 2016;374:2357-66. [PMID: 27075832] doi:10.1056/NEJMsa1600142
3. **McWilliams JM.** Changes in Medicare Shared Savings Program savings from 2013 to 2014 [Letter]. JAMA. 2016;316:1711-3. [PMID: 27612292] doi:10.1001/jama.2016.12049
4. **Colla CH, Lewis VA, Kao LS, O'Malley AJ, Chang CH, Fisher ES.** Association between Medicare accountable care organization implementation and spending among clinically vulnerable beneficiaries. JAMA Intern Med. 2016;176:1167-75. [PMID: 27322485] doi:10.1001/jamainternmed.2016.2827
5. **Schur CL, Sutton JP.** Physicians in Medicare ACOs offer mixed views of model for health care cost and quality. Health Aff (Millwood). 2017;36:649-54. [PMID: 28373330] doi:10.1377/hlthaff.2016.1427
6. **Markovitz AA, Ramsay PP, Shortell SM, Ryan AM.** Financial incentives and physician practice participation in Medicare's value-based reforms. Health Serv Res. 2018;53 Suppl 1:3052-69. [PMID: 28748535] doi:10.1111/1475-6773.12743
7. **Lewis VA, Colla CH, Carluzzo KL, Kler SE, Fisher ES.** Accountable care organizations in the United States: market and demographic factors associated with formation. Health Serv Res. 2013;48:1840-58. [PMID: 24117222] doi:10.1111/1475-6773.12102
8. **Hsu J, Vogeli C, Price M, Brand R, Chernew ME, Mohta N, et al.** Substantial physician turnover and beneficiary 'churn' in a large Medicare Pioneer ACO. Health Aff (Millwood). 2017;36:640-8. [PMID: 28373329] doi:10.1377/hlthaff.2016.1107
9. **Markovitz AA, Hollingsworth JM, Ayanian JZ, Norton EC, Moloci NM, Yan PL, et al.** Risk adjustment in Medicare ACO program deters coding increases but may lead ACOs to drop high-risk beneficiaries. Health Aff (Millwood). 2019;38:253-61. [PMID: 30715995] doi:10.1377/hlthaff.2018.05407
10. **Centers for Medicare & Medicaid Services.** Shared savings and losses and assignment methodology: specifications (version 3). 2014. Accessed at www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/sharedsavingsprogram/Downloads/Shared-Savings-Losses-Assignment-Spec.pdf on 6 January 2019.

11. **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;313:2152-61. [PMID: 25938875] doi:10.1001/jama.2015.4930
12. **Gottlieb DJ, Zhou W, Song Y, Andrews KG, Skinner JS, Sutherland JM.** Prices don't drive regional Medicare spending variations. Health Aff (Millwood). 2010;29:537-43. [PMID: 20110290] doi:10.1377/hlthaff.2009.0609
13. **Agency for Healthcare Research and Quality.** Prevention quality indicators technical specifications - version 6.0. 2016. Accessed at www.qualityindicators.ahrq.gov/modules/PQI_TechSpec.aspx on 6 January 2019.
14. **RTI International.** Accountable care organization 2014 program analysis quality performance standards narrative measure specifications. 2014. Accessed at www.healthreform.ct.gov/ohri/lib/ohri/work_groups/quality/2014-09-03/aco_narrative_measures_specs_1.pdf on 6 January 2019
15. **Prasad V, Jena AB.** Prespecified falsification end points: can they validate true observational associations? JAMA. 2013;309:241-2. [PMID: 23321761] doi:10.1001/jama.2012.96867
16. **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;308:1015-23. [PMID: 22968890] doi:10.1001/2012.jama.10812
17. **Fisher ES, Wennberg JE, Stukel TA, Sharp SM.** Hospital readmission rates for cohorts of Medicare beneficiaries in Boston and New Haven. N Engl J Med. 1994;331:989-95. [PMID: 8084356]
18. **Wennberg JE, Freeman JL, Shelton RM, Bubolz TA.** Hospital use and mortality among Medicare beneficiaries in Boston and New Haven. N Engl J Med. 1989;321:1168-73. [PMID: 2677726]
19. **McClellan M, McNeil BJ, Newhouse JP.** Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables. JAMA. 1994;272:859-66. [PMID: 8078163]
20. **Valley TS, Sjoding MW, Ryan AM, Iwashyna TJ, Cooke CR.** Association of intensive care unit admission with mortality among older patients with pneumonia. JAMA. 2015;314:1272-9. [PMID: 26393850] doi:10.1001/jama.2015.11068
21. **Bekelis K, Marth NJ, Wong K, Zhou W, Birkmeyer JD, Skinner J.** Primary stroke center hospitalization for elderly patients with stroke: implications for case fatality and travel times. JAMA Intern Med. 2016;176:1361-8. [PMID: 27455403] doi:10.1001/jamainternmed.2016.3919
22. **Crooks VA, Schuurman N.** Interpreting the results of a modified gravity model: examining access to primary health care physicians in five Canadian provinces and territories. BMC Health Serv Res. 2012;12:230. [PMID: 22852816] doi:10.1186/1472-6963-12-230
23. **Davidson R, MacKinnon JG.** Estimation and Inference in Econometrics. New York: Oxford Univ Pr; 1993.
24. **Luo W, Wang F.** Measures of spatial accessibility to health care in a GIS environment: synthesis and a case study in the Chicago region. Environ Plan B Urban Anal City Sci. 2003;30:865-84.
25. **McGrail MR, Humphreys JS.** Measuring spatial accessibility to primary care in rural areas: improving the effectiveness of the two-step floating catchment area method. Appl Geogr. 2009;29:533-41.
26. **WWAMI Rural Health Research Center.** Rural Urban Commuting Area Data: ZIP Code RUCA Approximation. Version 2.0. Seattle: WWAMI Rural Health Research Center; 2006.
27. **Winblad U, Mor V, McHugh JP, Rahman M.** ACO-affiliated hospitals reduced rehospitalizations from skilled nursing facilities faster than other hospitals. Health Aff (Millwood). 2017;36:67-73. [PMID: 28069848] doi:10.1377/hlthaff.2016.0759
28. **McWilliams JM, Chernew ME, Landon BE, Schwartz AL.** Performance differences in year 1 of Pioneer Accountable Care Organizations. N Engl J Med. 2015;372:1927-36. [PMID: 25875195] doi:10.1056/NEJMsa1414929
29. **United States Census Bureau.** American FactFinder. 2017. Accessed at https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml on 6 January 2019.
30. **McWilliams JM, Gilstrap LG, Stevenson DG, Chernew ME, Huskamp HA, Grabowski DC.** Changes in postacute care in the Medicare Shared Savings Program. JAMA Intern Med. 2017;177:518-26. [PMID: 28192556] doi:10.1001/jamainternmed.2016.9115
31. **Baiocchi M, Cheng J, Small DS.** Instrumental variable methods for causal inference. Stat Med. 2014;33:2297-340. [PMID: 24599889] doi:10.1002/sim.6128
32. **Staiger D, Stock J.** Instrumental variables regression with weak instruments. Econometrica. 1997;65:557-86.
33. **Baum CF, Schaffer ME, Stillman S.** Instrumental variables and GMM: estimation and testing. Stata J. 2003;3:1-31.
34. **Clough JD, McClellan M.** Implementing MACRA: implications for physicians and for physician leadership. JAMA. 2016;315:2397-8. [PMID: 27213914] doi:10.1001/jama.2016.7041

Current Author Addresses: Mr. Markovitz: UM Institute for Health Policy and Innovation, Center for Evaluating Health Reform, North Campus Research Complex, 2800 Plymouth Road, Building 16, 016-100N-28, Ann Arbor, MI 48109.
Dr. Hollingsworth: University of Michigan Medical School, Urology, North Campus Research Complex, 2800 Plymouth Road, Building 16, 016-112W, Ann Arbor, MI 48109-2800.
Dr. Ayanian: University of Michigan Medical School, UM Institute for Health Policy and Innovation, North Campus Research Complex, 2800 Plymouth Road, Building 16, 440E, Ann Arbor, MI 48109-2800.
Dr. Norton: University of Michigan School of Public Health, M3108 SPH II, 1415 Washington Heights, Ann Arbor, MI 48109-2029.
Ms. Yan: University of Michigan Medical School, Urology, North Campus Research Complex, 2800 Plymouth Road, Building 16, 016-100S-14, Ann Arbor, MI 48109-2800.
Dr. Ryan: University of Michigan School of Public Health, M3124 SPH II, 1415 Washington Heights, Ann Arbor, MI 48109-2029.

Author Contributions: Conception and design: A.A. Markovitz, J.M. Hollingsworth, J.Z. Ayanian, E.C. Norton, A.M. Ryan.
Analysis and interpretation of the data: A.A. Markovitz, J.M. Hollingsworth, J.Z. Ayanian, E.C. Norton, A.M. Ryan.
Drafting of the article: A.A. Markovitz.
Critical revision of the article for important intellectual content: A.A. Markovitz, J.M. Hollingsworth, J.Z. Ayanian, A.M. Ryan.
Final approval of the article: A.A. Markovitz, J.M. Hollingsworth, J.Z. Ayanian, E.C. Norton, P.L. Yan, A.M. Ryan.
Statistical expertise: A.A. Markovitz, E.C. Norton, A.M. Ryan.
Obtaining of funding: A.A. Markovitz, J.M. Hollingsworth.
Administrative, technical, or logistic support: A.M. Ryan.
Collection and assembly of data: J.M. Hollingsworth, P.L. Yan.