

HARVARD MEDICAL SCHOOL  
DEPARTMENT OF HEALTH CARE POLICY

180 LONGWOOD AVENUE  
BOSTON, MA 02115-5899



MAIN OFFICE: (617) 432-3333  
FAX: (617) 432-0173

September 18, 2018

Seema Verma, MPH  
Administrator, Centers for Medicare & Medicaid Services  
U.S. Department of Health and Human Services  
Hubert H. Humphrey Building, 200 Independence Avenue, S.W.  
Washington, DC 20201

RE: [CMS-1701-P] Medicare Program; Medicare Shared Savings Program; Accountable Care Organizations—Pathways to Success; proposed rule

Dear Administrator Verma:

We appreciate the Centers for Medicare and Medicaid Services' (CMS) efforts to improve the Medicare Shared Savings Program (MSSP) and agree with the fundamental goal of the proposed rule—to strengthen incentives for accountable care organizations (ACOs) to reduce wasteful health care spending. As academic researchers, we have published numerous studies assessing the effects of Medicare ACO programs on spending and quality and have written extensively about the weak incentives in the MSSP and the options for reform.<sup>1</sup> We believe that health policy should be evidence-based to the extent possible and should anticipate the dynamic consequences of changes in incentives. Although we agree that stronger incentives for ACOs to reduce Medicare spending could accelerate the early success of the MSSP, we are concerned that several key provisions in the proposed rule are at odds with existing evidence and could lead, conversely, to higher Medicare spending when implemented in the context of a voluntary program.

---

<sup>1</sup> McWilliams, J. M., B. E. Landon, et al. (2014). "Changes in Patients' Experiences in Medicare Accountable Care Organizations." *N Engl J Med* 371(18): 1715-1724. McWilliams, J. M., M. E. Chernew, et al. (2015). "Performance Differences in Year 1 of Pioneer Accountable Care Organizations." *N Engl J Med* 372(20): 1927-1936. McWilliams, J. M., L. A. Hatfield, et al. (2016). "Early Performance of Accountable Care Organizations in Medicare." *N Engl J Med* 374(24): 2357-2366. McWilliams, J. M., L. A. Hatfield, et al. (2018). "Medicare Spending after 3 Years of the Medicare Shared Savings Program." *N Engl J Med*. Advance online publication. <https://www.nejm.org/doi/full/10.1056/NEJMsa1803388>. McWilliams, J. M. (2016). "Changes in Medicare Shared Savings Program Savings from 2013 to 2014." *JAMA* 316(16): 1711-1713. McWilliams, J. M., M. E. Chernew, et al. (2014). "Outpatient Care Patterns and Organizational Accountability in Medicare." *JAMA Intern Med* 174(6): 938-945. Schwartz, A. L., M. E. Chernew, et al. (2015). "Changes in Low-Value Services in Year 1 of the Medicare Pioneer Accountable Care Organization Program." *JAMA Intern Med*. 175(11): 1815-1825. McWilliams, J. M., M. E. Chernew, et al. (2017). "Medicare ACO Program Savings Not Tied To Preventable Hospitalizations or Concentrated Among High-Risk Patients." *Health Affairs* 36(12): 2085-2093. Douven, R., T. G. McGuire, et al. (2015). "Avoiding Unintended Incentives in ACO Payment Models." *Health Affairs* 34(1): 143-149. Rose, S., A. M. Zaslavsky, et al. (2016). "Variation in Accountable Care Organization Spending and Sensitivity to Risk Adjustment: Implications for Benchmarking." *Health Affairs*. 35(3): 440-448. Chernew, M. E., T.M. McGuire, J. M. McWilliams. Refining the ACO Program: Issues and Options [Internet]. Boston (MA): Harvard Medical School; 2014 Nov 7 [cited 2018 Sept 10]. (Healthcare Markets and Regulations Lab Working Paper). Neprash, H. T., M. E. Chernew, et al. (2017). "Little Evidence Exists To Support The Expectation That Providers Would Consolidate To Enter New Payment Models." *Health Affairs* 36(2): 346-354. Barnett, M. L. and J. M. McWilliams (2018). "Changes in specialty care use and leakage in Medicare accountable care organizations." *Am J Manag Care* 24(5): e141-e149.

In principle, we agree with a payment model that holds providers in the same market to the same standard of efficiency after fully accounting for patient factors. Implementing such a model too quickly, however, will likely have harmful effects on participation that act to increase rather than decrease Medicare spending. In short, the proposed convergence of ACO benchmarks within markets could cause the voluntary MSSP to unravel, particularly when additional risk-bearing is imposed. Thus, we believe that the primary goal of the MSSP *at this stage* should be to reduce excessive spending so as to cause convergence in spending between providers and allow future consideration of a longer-term model in which payments might be set regionally as in Medicare Advantage. As long as the MSSP is voluntary and there is wide variation in spending, this goal can only be attained by strongly encouraging the participation of providers serving patients who receive less efficient care and establishing strong incentives for those providers to reduce wasteful spending. Below, we summarize and then detail our key concerns, their empirical and conceptual bases, and our recommended changes to the rule. We sincerely hope our analyses and comments are helpful in your deliberations and would be happy to discuss them further.

## **SUMMARY**

Our chief concern is that the combination of requiring risk for spending in excess of benchmarks (downside risk) and earlier blending of historical benchmarks with average regional spending via the regional adjustments will introduce even stronger incentives for ACOs with high spending for their patients to exit the MSSP and for providers with already low spending to remain in or join the MSSP. Such systematic selection has already occurred and will grow worse as regionalization of benchmarks and downside risk are imposed. There is strong evidence of strategic participation decisions by ACOs already (summarized below). If the MSSP is to cause convergence in spending by encouraging ACOs with high spending to reduce wasteful care, it needs to retain and attract providers with high spending and give them strong incentives to lower it. Although downside risk and regionalized benchmarks strengthen incentives to save, only providers who participate will be exposed to the stronger incentives. As ACOs with high spending exit and those with low spending enter and continue in the MSSP, ACOs will appear to be performing better against their benchmarks (ostensibly greater savings per program metrics) but the true impact of the program will be a substantial increase in Medicare spending in the form of subsidies to already more efficient ACOs. To the extent that risk-adjustment is incomplete, regionalizing benchmarks will also cause unmerited financial transfers from providers serving healthy and socially advantaged patients to those serving sicker and disadvantaged patients.

In addition to causing significant participation losses of key ACOs, imposing downside risk may do little to strengthen incentives. Remaining ACOs will tend to have low spending for their region and will thus be exposed to little additional risk because their spending is already below a regionally adjusted benchmark. In addition, many ACOs that are at risk of losses will seek reinsurance or partnerships with health management firms to mitigate risk (also weakening upside incentives as such firms take a cut of savings). The involvement of reinsurance and management firms will also add to the administrative costs of the program, eroding a key cost advantage of the ACO model over Medicare Advantage. It is also important to recognize that the initial structure of the program (one-sided contracts with benchmarks based on historical spending) has produced meaningful gains,

particularly among physician groups and ACOs with high initial spending.<sup>2,3,4,5</sup> While incentives have been very weak and should be strengthened, we believe that it is important to refine the MSSP in ways that build on those early gains rather than risk reversion and a squandering of providers' demonstrated willingness to try to reign in excessive spending.

Therefore, while we agree with the goal of strengthening incentives to save, we recommend a direction that differs from the proposed rule in key respects. Specifically, to encourage participation and foster convergence in spending while the MSSP remains a voluntary program, we recommend eliminating regional benchmark adjustments, eliminating the rebasing of benchmarks that has greatly diminished ACO incentives to ever save, and increasing the shared savings rate according to an ACO's spending relative to their regional average. Given the evidence that physician group ACOs have generated greater savings than larger hospital-integrated organizations, we agree with the spirit of proposed provisions that favor participation by low-revenue groups but would go even further to do so. Specifically, we recommend that requirements to assume downside risk be applied only to ACOs with high revenues on par with those generated by health systems providing inpatient, outpatient, and specialty care (i.e., accounting for a high proportion of Medicare spending for their patients). Lower-revenue ACOs should still be able to elect downside risk in exchange for favorable regulatory treatment and higher shared-savings rates. Program results and design should be revisited regularly, with monitoring of within-market spending variation to guide consideration of a longer-term payment model at a future time. Such an approach would establish strong incentives for all ACOs to lower wasteful spending, regardless of initial efficiency, by allowing ACOs time to recoup up-front investments and by eliminating penalties (in the form of lower benchmarks) for achieving savings. Such an approach also would afford the time necessary to modernize risk-adjustment methods to address "upcoding" and unintended consequences for disparities.

In addition to these core recommendations, we also offer suggestions for improving the methods for risk adjustment, trending benchmark updates, addressing compositional changes in ACO provider participants, and determining ACO organizational revenue.

In summary, we believe a conclusion that the MSSP has been a failure after 6 years of operation is inappropriate. The incentives for participating ACOs to lower spending have always been weak—the program can't have failed if there was never much opportunity for success. Moreover, the MSSP has generated meaningful savings from physician groups, for whom the incentives have been stronger. Ideally at the outset, the program would have had sufficiently strong incentives to allow more definitive judgment of its success. Delivery system reform is a slow process, and the voluntary nature of the MSSP limits the options for strengthening incentives, but incentives to save in a voluntary program can be strong without discouraging participation. For these reasons, we believe that the proposed changes to ACO benchmarking and risk-sharing are premature and counterproductive and that other available strategies should be pursued to accelerate progress.

---

<sup>2</sup> McWilliams, J. M., M. E. Chernew, et al. (2015). *N Engl J Med*.

<sup>3</sup> McWilliams, J. M., L. A. Hatfield, et al. (2016). *N Engl J Med*.

<sup>4</sup> McWilliams, J. M., L. A. Hatfield, et al. (2018). *N Engl J Med*.

<sup>5</sup> McWilliams, J. M. (2016). *JAMA*.

## Key Points

- The MSSP has generated net savings to Medicare
- Benchmarks are not valid counterfactuals
- Regional blending of benchmarks will cause selective exit from a voluntary MSSP; there is evidence already of cost-increasing selection in response to regional adjustments
- Incentives in the MSSP have been weak for reasons other than lack of downside risk
- There are major drawbacks to imposing downside risk
- Claims that downside risk has led to greater savings for Medicare are unfounded
- It is not clear that downside risk will selectively deter program abuse; participation to engage in anticompetitive behavior may be more profitable than earning shared savings
- Participation losses will be costly to Medicare; ACOs that did not earn bonuses and exited produced net savings to Medicare before exiting (n.b., benchmarks ≠ counterfactuals)
- Risk adjustment methods need to be modernized
- Determination of growth rates for benchmark updates should exclude an ACO's assigned population and can use standard statistical approaches to handle small sample sizes

## Key Recommendations

1. Because benchmark-based calculations are not informative for assessing program savings, base assessments of program impact solely on studies that are well designed to estimate it. Benchmarks need not be redesigned to be valid counterfactuals—their purpose is to set incentives—but they should not be used for program evaluation.
2. Set the shared-savings rate at a minimum of 50% for all ACOs in all performance years and tracks.
3. Eliminate rebasing of benchmarks that links an ACO's new benchmark to its spending changes in the prior period. This link should be severed completely. Instead, an ACO's original baseline should be updated annually according to a growth rate that is unrelated to its prior performance, adjusting for changes in the ACO's participating providers.
4. Eliminate regional adjustments that blend ACO historical benchmarks with regional spending. Replace with shared savings rates that rise continuously with lower ACO spending relative to its region, starting at 50% for ACOs with the highest spending for their region and rising to 80% for ACOs with the lowest spending for their region.
5. Eliminate downside risk requirements for lower-revenue ACOs by retaining Track 1, eliminating the BASIC track, and allowing voluntary participation in Track 1+. Replace the "clinician-led" basis for determining loss limits in Track 1+ with the proposed revenue-based determination (revenue/total spending for ACO population).
6. Continue to require high-revenue ACOs to assume downside risk commensurate with the ENHANCED track but set the threshold for defining high revenue at a higher level corresponding to an integrated health system. Also require downside risk of very large physician groups with market power. Rather than rely solely on participant lists, use additional data on full organizational structure to determine organization-wide revenue for core practices responsible for attribution.
7. Eliminate cap on *downward* benchmark adjustments that result from *decreases* in risk scores to mitigate what would otherwise be strong incentives to engage in favorable risk selection. Use ACO CAHPS data to help distinguish between changes in coding intensity and changes in population health status. Should benchmarks be converged to a regional

average in the future (not in the near term per Recommendation 4), implement additional payments or benchmark adjustments so that ACOs are not penalized for serving higher-risk patients.

8. Remove an ACO's attributed population from benchmark update determinations to preserve incentives for ACOs to lower spending. Use multilevel statistical modeling approaches to estimate ACO-specific update factors that reflect regional and national trends, with greater weight placed on national trends for ACO's that serve higher proportions of patients in their service area. After a phase of program expansion, switch to using a desirable preset rate of benchmark growth applied to all ACOs.

### **THE MSSP HAS GENERATED NET SAVINGS TO MEDICARE**

The use of benchmarks to assess the impact of the MSSP on Medicare spending has caused much confusion about whether the MSSP has produced savings to Medicare net of bonuses and the extent of those savings. Program benchmarks are important because they establish the incentives for ACOs to participate and lower spending. But understanding ACO responses to those incentives requires rigorous evaluation in which spending for ACO-attributed patients is compared with spending that would be expected to have occurred in the absence of ACO participation in the MSSP (the counterfactual). Benchmarks are not valid counterfactuals for a variety of reasons, several of which have caused benchmark-based assessments to underestimate savings in the MSSP (at least through 2016). First, until 2017, benchmarks were updated with the use of national spending growth rates (in absolute dollars/year). Because Medicare spending growth varies geographically, this feature of benchmark setting meant that an ACO's benchmark differed systematically from its counterfactual (expected) spending in almost all cases. The direction of the bias has differed for ACOs in areas of faster vs. slower spending growth, and the average effect has been to significantly underestimate expected spending because the national growth rate (updated factor) was determined among all FFS Medicare beneficiaries, including those with no primary care services to support attribution. Beneficiaries with no qualifying services, on average, have much lower spending and thus lower spending growth rates in absolute terms. For example, we estimate that Medicare spending growth from 2013 to 2016 was \$140/beneficiary slower among *all* beneficiaries than among those with at least one office visit with a primary care physician. Based on this difference, the national growth rate used to set benchmarks underestimated spending reductions achieved by ACOs for their primary care patients from 2013-2016 by approximately 1.3 percentage points.

Second, any spending reductions achieved by ACOs contribute to a slowing in national spending growth, causing benchmarks to fall. For example, based on our most recent analysis of MSSP savings,<sup>6</sup> spending reductions achieved by the 2012-2014 entry cohorts of ACOs caused their benchmarks in 2015 to be approximately 0.3 percentage points lower than if they did not lower spending. Third, rebasing of ACO benchmarks at the outset of second agreement periods that lowers benchmarks according to the extent of spending reductions achieved in the first period causes benchmarks in the second period to diverge even further below counterfactual spending, causing further underestimation of savings. This occurred for the 2012-2013 entry cohorts when they began their second agreement

---

<sup>6</sup> McWilliams, J. M., L. A. Hatfield, et al. (2018). N Engl J Med.

periods in 2016. Other features of benchmark setting have had less predictable effects but also could have contributed to underestimation of savings. For example, the approach to risk adjustment has appropriately tried to address other program goals, such as mitigating incentives to increase diagnosis coding intensity, but in the process may have not accurately reflected changes in the health risk of patients served by ACOs.

For these reasons, analyses using ACO benchmarks to quantify MSSP savings are uninformative and should be ignored when assessing program savings. Benchmarks are of obvious importance in establishing incentives and determining performance bonuses but are not valid counterfactuals and yield highly misleading results when used to assess program impact. For example, whereas benchmark comparisons would suggest that the MSSP caused cumulative net losses to Medicare of \$266M in 2014 and 2015 (including shared-savings bonuses), our evaluations, which use local concurrent spending changes to establish counterfactual spending, more accurately quantify the impact of the program as generating \$431M in cumulative savings to Medicare over those two years after accounting for bonuses — a \$697M difference from benchmark-based savings estimates. These greater savings are also consistent with market-level analyses conducted by government analysts and other researchers (including those summarized in the Regulatory Impact Analysis of the proposed rule). The review of the evidence in the Regulatory Impact Analysis is at odds with characterizations of one-sided contracts in the MSSP (e.g., on p.41787) as an unaffordable option that increases Medicare spending.

Although the direct savings we estimated are modest, the total savings to Medicare are greater because of spillover effects.<sup>7</sup> Changes in care for patients attributed to ACOs are likely to affect Medicare beneficiaries served by ACOs but not attributed to one.<sup>8</sup> Moreover, reductions in FFS Medicare spending achieved by ACOs in turn lower payment rates to Medicare Advantage plans. Particularly because of the large size of the growing Medicare Advantage program, these spillovers are substantial, approximately doubling the net savings to Medicare from the MSSP based on our previously published calculations.<sup>9</sup>

Moreover, the savings have varied with the strength of incentives to which ACOs are exposed, suggesting that refinements that strengthen incentives without compromising participation would lead to greater savings. In particular, we have consistently found that savings are significantly greater among physician-group ACOs than hospital-integrated ACOs.<sup>10,11,12,13</sup> Because physician groups provide less of the care spectrum, they have much

---

<sup>7</sup> McWilliams, J. (2016). "Savings from ACOs—Building on Early Success." *Annals of Internal Medicine* 165(12): 873-875.

<sup>8</sup> McWilliams, J., B. E. Landon, et al. (2013). "Changes in Health Care Spending and Quality for Medicare Beneficiaries Associated with a Commercial ACO Contract." *JAMA* 310(8): 829-836.

<sup>9</sup> McWilliams, J. (2016). *Annals of Internal Medicine*.

<sup>10</sup> McWilliams, J. M., L. A. Hatfield, et al. (2016). *N Engl J Med*.

<sup>11</sup> McWilliams, J. M. (2016). *JAMA*.

<sup>12</sup> McWilliams, J. M., L. A. Hatfield, et al. (2018). *N Engl J Med*.

<sup>13</sup> "Our definition of a physician group ACO is based on ownership structure and thus differs from CMS's categorization of an ACO as "physician-led" if there are no hospitals listed in its participant list. Thus, medical groups that are partnering with a hospital but do not share an ownership structure with the hospital are considered independent physician groups according to our categorization. Similarly, a health system that includes only its physician groups in its ACO participant list is considered hospital-integrated by our definition, not an independent physician group."

stronger incentives to reduce spending. All else equal, ACOs have stronger incentives to lower spending on care they *do not* provide than care they *do* provide. When ACOs lower spending on care they provide, any resulting shared-savings bonuses are offset to some extent by foregone fee-for-service profits. Similarly, spillover effects on non-ACO patients (including commercially insured patients) have greater financial consequences for larger hospital-based ACOs. If a primary care-focused ACO implements systemic care redesigns to reduce excessive inpatient or specialty care for all of its patients (not just those covered by its MSSP contract), it does not incur losses in revenue for non-ACO patients. Thus, physician groups have much stronger incentives as ACOs to lower spending and to implement more systemic changes that affect care for all patients. Not only have savings been convincingly higher for physician group ACOs but they have also been growing over time consistently in each MSSP entry cohort.<sup>12</sup> In our analyses, we have also confirmed that the greater savings estimated for physician groups are not due to those ACOs being located in higher-spending or higher-use regions or having higher spending for their regions (we obtain similar results when adjusting for those factors). *Because physician groups effectively have a higher shared-savings rate (because they incur lower offsetting costs when lowering spending), these findings foreshadow the additional savings that may be realized if shared savings rates are increased.*

Savings have varied with strength of incentives in other ways, too. For example, savings have been particularly pronounced in post-acute care, a leading source of waste in Medicare which most ACOs do not provide and thus almost universally have strong incentives to limit. In addition, assuming the costs of reducing spending are lower when spending is higher (where there is more waste it is easier to cut), savings should be greater for ACOs with higher baseline spending for their region or in higher-spending regions. Indeed, that has been observed as well.<sup>10</sup>

**Recommendation 1:** *Because benchmark-based calculations are not informative for assessing program savings, we urge CMS to base assessments of program impact solely on studies that are well designed to estimate it. Benchmarks need not be redesigned to be valid counterfactuals—their purpose is to set incentives—but they should not be used for program evaluation.*

### **CURRENT INCENTIVES WEAK FOR REASONS OTHER THAN LACK OF DOWNSIDE RISK**

There is near-universal agreement that incentives have been weak in the MSSP, but the discourse has focused almost exclusively on the lack of downside risk in MSSP contracts as the cause of weak incentives. The proposed rule intends to correct this perceived flaw by requiring currently participating high-revenue ACOs to assume significant downside risk immediately by entering the ENHANCED track, requiring new high-revenue participants to enter the ENHANCED track after 5 years, and requiring all new entrants and continuing low-revenue ACOs to assume some downside risk after 2 years of participation in the BASIC track. However, there are other major reasons why incentives to lower spending have been weak in the MSSP and thus other options for strengthening them.

First, shared savings rates have been low in the MSSP, allowing ACOs in Track 1 to keep no more than 50% of the difference between its expenditures and its benchmark. Because

imperfect quality scores reduce the shared savings rate, it has been even lower, averaging 44.2% in 2014 and 47.8% in 2017. Such shared-savings rates offer limited opportunities to profit relative to the status quo because ACOs incur substantial costs from being an ACO and attempting to earn shared savings.<sup>14</sup> These include the costs of ACO governance and quality reporting, efforts to lower spending and improve performance on quality measures, and (if those efforts are successful) any offsetting losses in profits as described above. Accordingly, the proposed reduction in the shared savings rate to 25-35% for ACOs in all levels of the BASIC track except Level E is likely to eliminate opportunities for ACOs to save on net, reduce ACO investments, and therefore reduce the savings to Medicare generated by ACOs and attract fewer ACOs into the program.

Second, the rebasing of the historical component of ACO benchmarks every 3 years to reflect an ACO's most recent 3 years of spending greatly diminishes ACOs' incentives to ever save because spending reductions now are penalized with lower benchmarks later.<sup>15</sup> To the extent that the costs of lowering spending are fixed (e.g., upfront investments with permanent effects on spending), rebasing reduces the time ACOs have to recoup those costs. To the extent that the costs of lowering spending are variable (must be incurred each time spending is reduced), expending effort to reduce wasteful services in a given year might contribute to shared savings for that performance year (if the minimum savings rate is exceeded) but will definitely cause the benchmark in the next agreement period to be lower and below the expected spending level under a status quo in which the effort is not expended again. Thus, in a contract with symmetric upside and downside risk, variable cost interventions (even if effective) pose a present-day net cost to ACOs.<sup>15</sup>

For example, let's say it costs a physician group ACO \$1000 in care manager salary to prevent one \$10,000 hospital admission through patient monitoring and timely engagement, diverting patients away from emergency departments and to the ACO's primary care practices when possible. For simplicity, assume 1-year agreement periods. The ACO will receive \$5,000 in shared savings for a net \$4,000 savings from such effort in the first agreement period, but its benchmark in the subsequent period will be lowered by \$10,000, posing a \$5,000 future loss (assuming 50% shared losses) and thus a net \$1000 present day cost from the care manager's effort if the ACO plans to stay in the program for a second period. The ACO can expend the \$1000 in care management again in the second period to avoid the penalty (it no longer produces a bonus from doing so because the benchmark is lowered), but as long as it participates in the program it must do so indefinitely, in every year of participation, to avoid the penalty (because spending will rise back to its previous level if the variable cost is not incurred). Thus, the ACO's incentives are either to never try to prevent admissions or to try to do so and then leave the program.<sup>16</sup>

The key point is that ACOs have had little incentive to ever lower spending because of the link between the spending reductions they achieve and subsequent benchmarks. One could argue that it is astonishing that the MSSP has produced any savings at all. Not surprisingly,

---

<sup>14</sup> Chernew, M. E., T.M. McGuire, J. M. McWilliams. Healthcare Markets and Regulations Lab Working Paper.

<sup>15</sup> Douven, R., T. G. McGuire, et al. (2015). Health Affairs.

<sup>16</sup> Douven, R., T. G. McGuire, et al. (2015). Health Affairs.



dropout in the 2012 and 2013 entry cohorts of ACOs was most marked when their benchmarks were set to be rebased at the start of their second agreement periods in 2016. Specifically, 44 of 191 ACOs (23%) left the program between 2015 and 2016 after annual exit rates averaging just 6.6% from 2013-2015 over the first agreement period.

**Recommendation 2:** *Set the shared-savings rate at a minimum of 50% for all ACOs in all performance years and tracks.*

**Recommendation 3:** *Eliminate rebasing of benchmarks that links an ACO's new benchmark to its spending changes in the prior period. This link should be completely severed. Instead, the ACO's original baseline should be updated annually according to a growth rate that is unrelated to its prior performance, adjusting for changes in the ACO's participating providers.*

Two concerns have been raised about the approach we are recommending, but we believe they have received too much weight in deliberations over MSSP rules to date and are addressable. First, without rebasing ACO benchmarks down to their more recent levels of spending, some have commented that this will allow ACOs with initially high spending to eventually earn windfall profits. We agree that some high-spending ACOs would be able to earn sizeable shared-savings bonuses eventually and believe that a principal goal of the MSSP should be to avail such opportunities. But shared savings are not equal to profits for the reasons we note above; they may be largely offset by the costs of lowering spending<sup>17</sup> and program participation. For many ACOs with initially high spending, particularly high-revenue ACOs that incur offsetting losses in FFS profits when restricting provision of care, the MSSP may merely offer an opportunity to mitigate those losses as they lower spending to levels of efficiency offered by other providers in their area. That may nevertheless be a goal of providers with high spending, particularly if they know with greater certainty the long-term direction of Medicare payment policy. But they will be unlikely to participate and pursue that goal if the costs are made too great by downside risk and regional blending. As we discuss below, concerns about an uneven playing field between ACOs with higher vs. lower spending in the same market can be addressed by increasing the shared-savings rate for more efficient ACOs. In addition, under both the current and proposed rules, benchmarks will be progressively reduced by the annual benchmark updates as the program matures and expands, because regional and national spending growth will progressively slow as ACOs increasingly lower Medicare spending. Thus, the annual updates will still apply downward pressure on ACO benchmarks, thereby limiting ACO profits. The key distinctions are that, with rebasing eliminated: 1) the downward pressure on benchmarks would be at a rate that appropriately lags behind the downward effect ACOs have on spending as opposed to at an equal rate (thereby encouraging participation); and 2) the downward pressure on a given ACO's benchmark would not be caused by its own efforts to lower spending (thereby strengthening incentives to save).

Second, concerns have been raised about the ability to adequately adjust for changes in ACO provider composition over longer periods of time. An inability to do so could allow ACOs to select more efficient practices (taxpayer identification numbers [TINs]) for

---

<sup>17</sup> Chernew, M. E., T.M. McGuire, J. M. McWilliams. Healthcare Markets and Regulations Lab Working Paper.

inclusion to earn shared savings without lowering spending. More generally, the clinicians and practices represented by a given TIN may change substantially because of mergers and acquisitions, and TINs may newly appear or cease to appear in claims. These challenges can be addressed with relatively straightforward approaches, several of which we have implemented in our evaluations of the MSSP.<sup>18,19</sup> When an ACO adds TINs to their participant list, the baseline period for those TINs could be established as the 3 years prior to the addition (i.e., different parts of the ACO would have different baseline periods). For TINs that exhibit large changes from baseline to performance years in the number of clinicians billing under them or that newly appear in claims, NPI-TIN billing relationships can be used to identify additional TINs in the original baseline period to assess baseline spending for a set of providers that is more consistent with an ACO's participant list in a given performance year. For TINs where provider consistency between performance and baseline years is tenuous and cannot be corrected with such methods, rebasing could be allowed for the portion of an ACO's benchmark accounted for by such TINs. Moreover, to the extent that the MSSP encourages ACOs to include more efficient providers in their contracts, that may help foster competition among providers to be efficient and attractive to ACOs. Thus, such behavior should not necessarily be viewed as wasteful gaming.

#### **PROBLEMS WITH REGIONAL BENCHMARK ADJUSTMENTS AND EVIDENCE OF COST-INCREASING PARTICIPATION DECISIONS IN THE 2014 COHORT**

The regional adjustment in the current benchmarking methodology, particularly when implemented in an ACO's first agreement period as proposed, is highly problematic in a voluntary MSSP, particularly if downside risk is imposed on participants. It will surely lead to exit by ACOs with spending above their region's average. These ACOs with high baseline spending have produced greater savings than ACOs with lower spending<sup>18,20</sup> and are critical to retain and attract if the MSSP is to ever cause sufficient within-market convergence in spending to facilitate a transition to a longer-term payment system in which providers bear risk and population-based benchmarks are similar for different providers in the same market (accounting for patient factors).

There is evidence already of ACOs with high spending for their region selectively exiting the program in response to the prospect of regional benchmark adjustments. As noted in the proposed rule, the 25-35% blending of ACOs' historical benchmarks with regional average spending led to substantial swings in benchmarks in 2017 for ACOs that started in 2014. This was to be expected from the documented wide variation ACOs' risk-adjusted spending relative to their service areas.<sup>21</sup> As described in the table below, introduction of the regional adjustment in 2017 was associated with 21% dropout in the 2014 cohort and more than a quadrupling of the mean "savings rate" among remaining participants (from 0.5% to 2.2%). The "savings rate" indicates where an ACO's spending was relative to its benchmark, with positive values indicating that spending was below the benchmark. To distinguish this quantity from estimates of actual ACO savings estimated by rigorous evaluations, we refer to this quantity as the "benchmark savings rate." The increase in the

<sup>18</sup> McWilliams, J. M., L. A. Hatfield, et al. (2016). *N Engl J Med*.

<sup>19</sup> McWilliams, J. M., L. A. Hatfield, et al. (2018). *N Engl J Med*.

<sup>20</sup> McWilliams, J. M., M. E. Chernew, et al. (2015). *N Engl J Med*.

<sup>21</sup> Rose, S., A. M. Zaslavsky, et al. (2016). *Health Affairs*.

benchmark savings rate in 2017 for the 2014 cohort (bolded in the table) was the most substantial increase observed in any entry cohort in any pair of years and thus unlikely due to improvement in ACO performance alone. The other year-to-year increases in benchmark savings rates were much more modest, including those associated with rebasing and extensive dropout in 2016 for the 2012-2013 cohorts. Of the remaining ACOs from the 2014 entry cohort, 81% had positive regional adjustments in 2017, meaning that their historical benchmarks were below average regional spending and thus that their participation was subsidized. *This pattern is highly suggestive of selective dropout of ACOs with spending above their regional average, causing the appearance of increased savings that would be more accurately characterized as subsidies to ACOs with lower spending.*

Entry year	2015 Performance		2016 Performance		2017 Performance	
	N	ACO-level Mean "Savings Rate" ([benchmark-spending]/benchmark), %	N	ACO-level Mean "Savings Rate" ([benchmark-spending]/benchmark), %	N	ACO-level Mean "Savings Rate" ([benchmark-spending]/benchmark), %
<b>2012</b>	100	1.4	73	1.8	63	2.6
<b>2013</b>	91	1.6	74	2.3	62	2.2
<b>2014</b>	112	-0.2	<b>100</b>	<b>0.5</b>	<b>79</b>	<b>2.2</b>
<b>2015</b>	89	-0.1	85	0.3	76	0.7
<b>2016</b>	-	-	100	0.2	96	0.9

Source: <https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/SSPACO/index.html>

From our analyses of Medicare claims data, we confirmed that ACOs in the 2014 cohort that exited the program in 2017 had substantially higher spending for their region than those who remained in the program. Specifically, we fit a model of total spending per beneficiary in 2014-15 as a function of county and year fixed effects and fixed effects for ACOs in the 2014 cohort. The latter set of fixed effects estimated the mean spending difference between each ACO in the 2014 cohort and its service area. *This spending deviation was, on average, \$787/beneficiary higher for ACOs dropping out than for those remaining in the MSSP.*

This pattern of cost-increasing selective participation will only grow worse if ACOs are required to assume downside risk. Even the limited amount of downside risk in the BASIC track (8% of Part A and B revenue) will easily exceed the bonus for qualifying as an Advanced APM (5% of Part B revenue) for many ACOs with high spending for their region, posing certain losses and likely making the MSSP a less appealing option than the status quo under the MIPS in the near term (particularly given the additional administrative and reporting costs borne by ACOs). This pattern will also grow worse in second and third agreement periods as the regional weight grows to 50%. *Thus, there is a significant risk the MSSP will degenerate into a program that is viable only for providers that are already more efficient for their region or serve patients who are healthier and lower-risk in ways not captured by the HCC score.* Capping the regional weight at 50% and capping the maximum adjustment at 5% of average total Medicare expenditures per beneficiary will still allow large swings in benchmarks and will do little to mitigate this risk of undeserved subsidies and exacerbated disparities.

We understand that the costs of lowering spending are likely greater for ACOs with lower spending (less “low-hanging fruit”) and that ACOs already providing more efficient care have voiced concerns that basing benchmarks solely on ACOs’ historical spending creates unfairly greater opportunities for providers who have been historically less efficient. *Providers who are more efficient than others, however, are not necessarily efficient.* For example, in our empirical work on low-value care in Medicare, we found that providers who provide the least low-value care still provide more low-value care than the difference between the most and least efficient providers.<sup>22,23</sup> Therefore, ACOs with low spending for their region should be able to reduce wasteful spending further, but likely require greater rewards to cover the higher costs of doing so. Therefore, to retain and attract ACOs with lower spending and to recognize their greater level of efficiency and the greater challenge they face in further reducing spending, we recommend replacing the regional adjustment with higher shared savings rates for ACOs with lower spending for their service area. This could be implemented in a graded fashion, with shared savings rates rising continuously from a minimum of 50% for ACOs with the highest spending for their region to 80% for ACOs with the lowest spending for their region, using the most recent 2 years to assess ACOs’ standing relative to their region. Relating the shared-savings rate to ACOs’ most recent level of spending will also strengthen incentives for ACOs to lower spending.

It is important to recognize that higher shared savings rates do not mean lesser savings for Medicare. First, higher shared savings rates should attract more providers into the MSSP. So even if Medicare receives a lower proportion of the savings for a given ACO, there should be more ACOs generating savings. Second, higher shared savings rates should elicit stronger efforts by ACOs to lower spending. To date, ACOs with spending below their regional average have not produced significant savings. Therefore, increasing the savings rate for an ACO with low spending for their region can only increase savings to Medicare from that ACO, on average. Third, because a higher shared savings rate does not apply to savings from spillover effects (e.g., lower Medicare Advantage payment rates), increasing the shared-savings rate will reduce the amount of savings to Medicare from a given ACO with a given performance level much less than the change in the percentage would indicate.

Finally, it is also important to bear in mind that most ACOs provide only some of the care spectrum (lower revenue ACOs). These ACOs tend to be primary care or multispecialty physician groups. Thus, much of the inefficient care delivered to an ACO’s population may not be provided by the ACO. Accordingly, such inefficient care offers opportunities for efficient physician groups to enter those communities and reduce spending as ACOs.

***Recommendation 4:*** *Eliminate regional adjustments that blend ACO historical benchmarks with regional spending. Replace with shared savings rates that rise continuously with lower ACO spending relative to its region, starting at 50% for ACOs with the highest spending for their region and rising to 80% for ACOs with the lowest spending for their region.*

---

<sup>22</sup> Schwartz, A. L., B. E. Landon, et al. (2014). "Measuring Low-value Care in Medicare." *JAMA Internal Medicine* 174(7): 1067-1076.

<sup>23</sup> Schwartz, A. L., A. M. Zaslavsky, et al. (2018). "Low-Value Service Use in Provider Organizations." *Health Serv Res* 53(1): 87-119.

## **DOWNSIDES OF DOWNSIDE RISK**

Not only can incentives in the MSSP be strengthened without imposing downside risk for spending in excess of benchmarks, but downside risk may ultimately strengthen incentives very little and will likely cause more problems than it solves, acting to diminish instead of increase savings to Medicare.

As a preliminary point, it is important to recognize that Track 1 is not riskless. As described above, all ACOs incur significant administrative and reporting costs to participate in the MSSP, and those that try to lower spending or improve quality incur additional costs. These costs pose a potential downside even if the risk-sharing is one-sided. ACOs lose money if they fail to earn sufficiently high shared-savings to offset the costs of participation and effort to achieve savings. As evidence that Track 1 is a downside risk track, the rate of dropout from the MSSP has been substantial, and ACOs that have dropped out are disproportionately those that did not receive shared-savings bonuses. Specifically by 2018, 40-45% of ACOs in the 2012-2013 entry cohort and 35% of the 2014 cohort had exited the MSSP. Benchmark savings rates were substantially lower among ACOs dropping out than among those staying in the program. In the 2012-2013 cohorts, for example, the mean benchmark savings rate among ACOs that have dropped out was 0.37% before they dropped out, compared with 1.44% among those continuing in the program. If Track 1 were riskless, ACOs not receiving bonuses would stay in the program. *Given that dropout has been substantial and driven largely by the limited performance-based risk already present in Track 1, we should expect that additional downside risk in the form of two-sided risk sharing will accelerate dropout and discourage participation.*

While one might argue that we may not want ACOs that cannot achieve savings to stay in the program, it is important to remember that *benchmarks are not counterfactuals*. Thus, ACOs that do not achieve shared-savings relative to their benchmarks may nevertheless contribute to actual savings to Medicare. Indeed, ACOs dropping out of the Pioneer model achieved similar spending reductions before dropping out compared with those that continued in the model (see McWilliams et al. NEJM 2015 and Nyweide et al. 2015). We reached a similar conclusion for the MSSP based on a similar analysis that we conducted after publishing our most recent evaluation (McWilliams et al. NEJM 2018). Specifically, we found that the ACOs in the 2012-2014 entry cohorts that dropped out of the MSSP by 2018 (approximately 40% of ACOs) achieved per-beneficiary spending reductions (gross savings) in 2015 that were, on average, about 20% smaller than the spending reductions achieved by ACOs that subsequently remained in the program. *But these spending reductions were statistically significant, and because shared-savings bonuses were smaller for drop outs, they accounted for \$113 million of the total \$145 million in net savings generated by the MSSP in 2015 overall (after accounting for shared-savings bonuses), and for \$109 million of the \$256 million in net savings generated by physician group ACOs in that year.* Thus, because benchmarks are not counterfactuals, they should not be used to determine which ACOs are contributing more or less savings to Medicare or which should be discouraged from participating. These findings suggest that dropout from the MSSP has been costly to Medicare and that it is worthwhile to retain ACOs that do not perform well against their benchmarks. In addition, some ACOs may take longer than one agreement period to successfully implement strategies to lower savings. For these reasons, we believe

that provisions that would further discourage participation by ACOs that do not perform well against their benchmarks (downside risk and regional adjustment of benchmarks) would be detrimental to the program.

Since the ENHANCED track is essentially the Pioneer model, the 50% dropout rate observed in the Pioneer model should also serve as a cautionary tale about the extent of voluntary participation that should be expected in the ENHANCED track. (Only 16 of the 32 Pioneer ACOs remained in two-sided models [Pioneer or Next Generation] by 2016). Limiting participation in the MSSP by high-revenue organizations, however, may prove to be sound policy if a goal of the MSSP is to foster competition by offering shared savings opportunities that appeal selectively to smaller organizations and physician groups.

As another point of evidence that is important for considering downside risk requirements, recent assertions that the downside risk tracks (2, 3, and now 1+) have caused greater savings are misleading and we believe incorrect. Some analysts, including the Medicare Payment Advisory Commission,<sup>24</sup> have observed that ACOs participating in Tracks 2 and 3 have higher benchmark savings rates and have concluded that downside risk elicits greater savings. However, only a small percentage of ACOs have participated in Track 2 or 3 (8% in 2018), and most of them transitioned into these tracks after experience in Track 1 (as of 2015, 99% of ACOs were in Track 1). Thus, these ACOs were able to select into Track 2 or 3 to take advantage of the higher shared-savings rates after ascertaining that their spending would continue to be below their benchmarks, whether because of preceding success in lowering spending or other factors that cause benchmarks to deviate from counterfactual spending. *Indeed, of the 32 ACOs that transitioned from Track 1 to Track 2 or 3 by 2018, the mean benchmark savings rate was 3.6% over the 3 years prior to changing tracks, vs. just 0.6% among other ACOs remaining in Track 1.* For the subgroup of ACOs that entered Track 2 or 3 in 2016 or 2017, the 1-2 years of downside risk exposure did not affect that pre-existing difference in benchmark savings rates, which remained similar in 2017 (3.5% for ACOs in tracks 2/3 vs. 1.3% for ACOs in Track 1), suggesting that entry into a track with downside risk contributed no additional savings. The mean benchmark savings rate in the 3 years prior to transitioning to Track 1+ was similarly high (2.3%) among ACOs leaving Track 1 for Track 1+ in 2018. The mean benchmark savings rate in 2017 among ACOs that entered Track 2 or 3 directly, without a prior agreement period in Track 1, was much lower (0.8%) than among ACOs selecting into downside risk tracks after experience in Track 1.

In addition, we know from evaluations of the Pioneer model<sup>25,26</sup> that spending reductions in that 2-sided model were no greater than those achieved in 1-sided contracts in the MSSP. Similarly, the evidence shows that the two-sided Next Generation (Next Gen) model has performed no better than Track 1 of the MSSP when accounting for duration of participation in an ACO program. Specifically, the recent evaluation of year 1 of the Next Gen model by NORC found gross annual savings of \$210/beneficiary in 2016, but 15 of the

---

<sup>24</sup> Report to the Congress: Medicare and the Health Care Delivery System (pp. 211-237, Rep.). (2018). Medical Payment Advisory Commission.

<sup>25</sup> McWilliams, J. M., M. E. Chernew, et al. (2015). *N Engl J Med*.

<sup>26</sup> Nyweide, D. J., W. Lee, et al. (2015). "Association of Pioneer Accountable Care Organizations vs Traditional Medicare Fee for Service with Spending, Utilization, and Patient Experience." *JAMA* **313**(21): 2152-2161.

18 ACOs participating in Next Gen had significant prior experience in the Pioneer model or MSSP, including 2 of the 4 ACOs that drove the bulk of the savings. These gross savings from Next Gen are no bigger than those achieved in the MSSP when appropriately compared with MSSP ACOs with multiple years of experience. For example, in our recently published evaluation of the MSSP, we found gross annual savings of \$213/beneficiary in 2015 for ACOs entering in 2012 or 2013 (slightly greater than the Next Gen estimate). *Thus, there is no evidence to date that downside risk has enhanced gross savings in Medicare ACO contracts, and thus no evidence that downside risk has elicited a stronger response from ACOs.*

Without eliciting stronger efforts by ACOs to lower spending, downside risk does allow CMS to recoup a portion of spending in excess of benchmarks. For example, the 2016 net savings per beneficiary in Next Gen (\$130/beneficiary in net savings to Medicare) were greater than the 2015 net savings per beneficiary in the 2012-2013 entry cohorts of MSSP ACOs (\$62/beneficiary) despite the similar gross savings. However, the net savings achieved by the 18 ACOs in Next Gen (most of which are large health systems) were smaller than the net savings generated by physician-group ACOs in the MSSP (\$180/beneficiary) and constituted much smaller aggregate net savings than produced by the more popular MSSP (\$62M from Next Gen vs. \$190M from all 2012-2013 entrants in the MSSP). This underscores the importance of allowing physician groups (lower revenue organizations) to participate in one-sided contracts and suggests that downside risk could be used by CMS to protect against losses from health systems, whose incentives in one-sided contracts are very weak. But we would note that *any losses recouped by CMS by requiring downside risk are likely to be at least partially, if not fully, offset by the increase in Advanced APM bonus payments for providers accepting downside risk.* In addition, because the Advanced APM bonus is structured as 5% of Part B revenue, it acts to weaken incentives for ACOs to reduce provision of wasteful Part B services (the size of the bonus is reduced when Part B revenue is reduced).

All else equal, we agree that introducing downside risk-sharing strengthens incentives for ACOs to lower spending. But all else will not remain equal if downside risk is imposed, particularly because the MSSP is voluntary and the proposed introduction of downside risk is coupled with regional blending of benchmarks. Since ACOs with spending above their regional average will disproportionately exit the program, most remaining ACOs will have spending below their regional average and thus benchmarks above their historical spending. For those ACOs, downside risk adds little additional incentive to lower spending below the status quo because the incentive to lower spending when spending is below the benchmark stems from the upside risk sharing provisions. The cost of maintaining the status quo for those ACOs is the foregone profits from not earning more shared savings, not penalties from downside risk. Thus, because of the anticipated participation losses, imposing downside risk may not significantly strengthen incentives for remaining ACOs.

Among ACOs that remain in the program and face a higher risk of losses relative to the status quo because their spending is at or above their region's average, many ACOs (particular smaller, lower revenue ACOs with lower financial reserves) will seek reinsurance, consolidate with other providers, or partner with a health management firm to pool risk with other ACOs in order to mitigate their downside risk. Even the limited amount of downside risk in the BASIC track (8% of revenue) poses potentially substantial

losses, particularly when sustained repeatedly over multiple years. Because many ACOs will seek to mitigate their risk of losses, the extent to which downside risk incentives will reach the levels of organization currently exposed to one-sided incentives is unclear. Since health management firms take a cut of any shared savings, requiring downside risk will also weaken upside risk incentives for ACOs that partner with such firms.

Moreover, provider responses to mitigate downside risk will act to erode a key advantage of the ACO model over the Medicare Advantage model – lower administrative costs from direct contracting between Medicare and providers. As more ACOs pay premiums for reinsurance premiums and fees to management firms (which may take the form of a cut of shared savings), the administrative costs of the program will grow.

Finally, a stated motivation in the proposed rule for imposing downside risk is to discourage abuse of the program by provider organizations that participate to exploit waivers from Stark and Anti-kickback law and engage in anticompetitive behavior. The proposed rule goes so far as to state, “Further, the presence of an ‘upside-only’ track may be encouraging consolidation in the marketplace, reducing competition and choice for Medicare FFS beneficiaries.” While this is certainly a valid concern, we are aware of no empirical evidence supporting this statement. In fact, research to date has found little evidence of greater consolidation among ACO participants as a result of the MSSP.<sup>27</sup> Consolidation had been occurring before the MSSP and has generally continued for other reasons. While there may have been some defensive consolidation among non-ACO providers<sup>27</sup> (e.g. to rebuff payer attempts to engage them in ACO risk contracts), that would not be addressed by changing risk-sharing provisions among MSSP participants. MSSP participation has also not been associated with detectable changes in referrals, as measured by leakage. Specifically, MSSP participation has not been associated with evidence of ACOs effectively steering outpatient specialty referrals internally to ACO specialists.<sup>28</sup> While this area remains one in need of further investigation, we do not believe it is appropriate to conclude that the MSSP has contributed to consolidation or anticompetitive control over referrals. Given that the MSSP model presents greater opportunities for physician groups (for the reasons described above) one could even argue that the MSSP could contribute to greater competition in provider markets, as long as it continues to favor lower-revenue organizational structures in its incentives.

Moreover, use of downside risk to deter gaming or anticompetitive behavior implicitly assumes that the financial gains from abusing the program are smaller than the gains from genuinely engaging in efforts to lower spending. However, the opposite is more likely to be the case. Attempts to lower spending are fraught with uncertainty, challenges, and offsetting costs, and have not been strongly rewarded in the MSSP because of rebasing and low shared savings rates, whereas engaging in anticompetitive behavior to enhance FFS profits is likely to be far more predictable and lucrative. A health system could participate in the MSSP to take advantage of waivers to maintain or increase referrals for inpatient and

---

<sup>27</sup> Neprash, H. T., M. E. Chernew, et al. (2017). Health Affairs.

<sup>28</sup> Barnett, M. L. and J. M. McWilliams (2018). "Changes in specialty care use and leakage in Medicare accountable care organizations." *Am J Manag Care* 24(5): e141-e149.



specialty care without affecting total Medicare spending (i.e., only causing a shift in care that favors the health system). Such behavior would increase provider revenue without resulting in penalties under downside risk. Thus, requiring downside risk is likely to selectively discourage ACOs joining the program for the right reasons, while ACOs joining the program for the wrong reasons may not be dissuaded by downside risk as much because the gains from regulatory relief may outweigh any potential losses. We would favor measures to identify abuses directly over blunt uses of downside risk requirements that risk having unintended consequences.

In summary, there is no evidence that downside risk has elicited greater savings in Medicare ACO programs, and imposing downside risk in a voluntary program with regionally adjusted benchmarks is likely to induce cost-increasing participation losses and higher administrative costs, without achieving the goals of stronger incentives and selectively deterring abuse.

**Recommendation 5:** *Eliminate downside risk requirements for lower-revenue ACOs by retaining Track 1, eliminating the BASIC track, and allowing voluntary participation in Track 1+. Replace the “clinician-led” basis for determining loss limits in Track 1+ with the proposed revenue-based determination (ACO revenue for Medicare patients/total Medicare spending for assigned population).*

**Recommendation 6:** *Continue to require high-revenue ACOs to assume downside risk commensurate with the ENHANCED track but set the threshold for defining high revenue at a higher level corresponding to an integrated health system. Also require downside risk of very large physician groups with market power. Rather than rely solely on participant lists, use additional data on full organizational structure to determine organization-wide revenue for core practices responsible for attribution.*

We make the last part of Recommendation 6 because we are concerned that ACOs, particularly high-revenue ACOs, may game revenue determinations if the ACO participant lists are used to determine ACO revenue, as proposed in the proposed rule. For example, a large health system could include only its provider organizations or primary care practices in order to be judged as a low-revenue ACO without significantly altering its attributed population. For example, a large health system may own several hospitals and multiple physician groups, each billing under distinct TINs. Or a health system may consist of a hospital system with a closely aligned provider organization billing under its own TIN. In such cases, which are common, the physician groups or provider organizations account for all or nearly all of the attribution but the hospitals and other facilities account for much of the revenue. Thus, an organization could include only its physician groups in its ACO contract and lower its downside risk without changing its assigned population or its de facto governance (because of the common ownership or mutual interests shared by the health system and the physician groups).

Relying on ACO participant lists to categorize ACOs as low-revenue or high-revenue also discourages partnerships between physician groups and hospitals through means other than mergers and acquisitions. Thus the proposed basis for revenue determinations could

discourage a key strategy taken by ACOs—steering volume to more efficient providers and formalizing preferences with inclusion in contracts. We note that many ACOs with hospitals in participant lists are actually independent physician groups, and we categorized them as such in our evaluations. These ACOs should be regarded as low-revenue to encourage their participation, but the proposed use of ACO participant lists could categorize them as high-revenue, incentivizing them to drop such partnerships. This could thwart healthy competition if such ACOs are able to get specialists and hospitals to compete on quality and efficiency for inclusion in ACO contracts.

Therefore, we recommend using additional data sources, such as IRS filings and PECOS data, to determine the full organizational structure of the core practices that account for the bulk of the assigned population, and to use all TINs and CCNs in the full organization to determine organizational revenue. We understand that PECOS data are used to validate TINs and CCNs in ACO participant lists, but these data can be additionally used by CMS to determine the full organization to which ACO participants belong. In other work funded by the AHRQ, we (a team of investigators at HMS and the NBER) have used these data sources to group independently billing TINs and CCNs into organizations or systems. Thus, although what we are suggesting would require some additional effort, we believe it would be feasible and would grow easier over time as such enhancements are made. We would be happy to discuss with CMS our experience in coalescing TINs and CCNs into organizations using these data sources.

Finally, we want to emphasize that we are recommending that *revenue determinations be based ideally on the full organizational structure of the core practices that determine attribution*. If an ACO includes physician groups that are part of a large health system, that ACO should be considered high-revenue whether or not the other parts of the health system are included as participants. If an ACO includes physician groups that are financially independent of hospitals, that ACO should be considered low-revenue (unless it is very specialty-heavy or has substantial market share), whether or not preferred hospital partners are included in the contracts. That way, consolidation in provider markets is discouraged with more downside-risk in ACO contracts, and partnerships or preferred networks that can support competition and do not cause commercial mark-ups are not discouraged. *The proposed use of only ACO participant lists for revenue determinations could constitute a lost opportunity for discouraging price-increasing consolidation.*

## **STRATEGIES FOR ADDRESSING INADEQUATE RISK ADJUSTMENT**

We agree with the proposed changes to account for increases in ACO population risk scores and to implement measures to limit rewards for increases in coding intensity. We do not, however, understand the rationale for limiting benchmark decreases to reflect decreases in risk scores. *Capping downward adjustments for changes in patient risk introduces strong incentives for ACOs to engage in favorable risk selection (i.e., cream-skimming).*

We believe there are several potential strategies for CMS to move beyond the HCC model, which is now 15 years old, to adjust more completely for patient risk, reduce incentives for providers to upcode and engage in risk selection, and minimize unwarranted penalties that could exacerbate disparities by depleting providers serving higher-risk patients of

resources available for quality improvement. In particular, the ACO CAHPS data, which include several measures of self-reported health and functional limitations, can be used to determine the extent to which increases in HCC scores reflect changes in coding vs. changes in health status. Discrepancies between the two could be used to limit benchmark increases due to coding increases in a more refined, ACO-specific manner. Such adjustments could be instituted after an initial grace period (e.g., 3 years) that would permit coding increases that are unrelated to health status changes since they might reflect efforts by under-coding ACOs to achieve accurate coding. Differences in spending predicted by HCC scores vs. CAHPS variables could also be used to identify ACOs that are undercoding relative to others at baseline and that should therefore be eligible for such a grace period.

Some of the clinical and social characteristics included in the CAHPS survey could also be used to further risk adjust ACO benchmarks, and this information has been found to explain a meaningful amount of variation in ACO spending after controlling for HCC score.<sup>29</sup> We would caution, however, against adjustments for time-varying health status variables, as adjusting for health status weakens incentives for ACOs to improve patient health.

*An advantage of using ACOs' historical spending as the basis for benchmarks is that historical spending accounts for all differences in patient population risk between ACOs, whether measurable or not. Blending benchmarks with regional average makes program success depend heavily on risk adjustment for observable patient factors.* In the context of inadequate risk adjustment, regionally adjusted benchmarks would cause unjustified financial transfers from ACOs serving sicker and poorer patients to those serving healthier and wealthier patients. Unfortunately, risk adjustment will always be inadequate because all differences between patient populations can never be fully accounted for, particularly with available data sources. *It is not a question of whether risk adjustment is inadequate but the extent of its inadequacy.*

Accordingly, in the future, should benchmarks be converged toward a regional average (again, we strongly recommend against regional adjustments at this juncture), we would urge CMS to implement additional measures to ensure that ACOs are not penalized for serving higher-risk patients, independent of their efficiency or quality of care. For example, an additional adjustment could take the form of a per-beneficiary care management fee that is higher for higher-risk patients, as in the Comprehensive Primary Care Plus program. To keep such a measure budget neutral and further encourage participation in the MSSP, a small discount could be applied to FFS rates paid to all providers and redistributed as care management fees to ACOs, with higher payments to ACOs serving higher-risk patients regardless of their performance. Alternatively, ACO benchmarks could simply be adjusted upward or downward by an additional amount (i.e., beyond adjustments for patient factors included in the risk-adjustment methodology) depending on whether their populations are higher- or lower-risk, respectively, than the regional average.

Finally, we are concerned that the proposed structure of the beneficiary incentive program could exacerbate favorable risk selection in the MSSP. Specifically, because an ACO must

---

<sup>29</sup> Rose, S., A. M. Zaslavsky, et al. (2016). Health Affairs.

bear the costs of the incentive payments and because an ACO must furnish an incentive payment to an eligible beneficiary each time the beneficiary receives a qualifying service (in accordance with section 1899(m)(2)(D)(iv) of the Bipartisan Budget Act), the costs to an ACO of serving higher-risk patients will be greater than the costs of serving lower-risk patients (since higher-risk patients use more services). For this reason, we believe it is important for beneficiary incentive programs to be optional (as proposed), expect few ACOs to establish one, and do not expect the opportunity to establish one will attract ACOs into downside risk tracks. We recommend close monitoring of ACOs that do implement beneficiary incentive programs.

In general, we are in favor of patients sharing in the savings generated by their ACO, whether in the form of direct payments, lower premiums, or in-kind benefits. In the future, once more convergence in spending has occurred and benchmarks reliably reward more efficient providers, such mechanisms for sharing savings with beneficiaries could help to make more efficient providers more attractive to beneficiaries and foster competition to achieve savings (to compete for patients), as long as the benefits to beneficiaries are greater when ACO savings are greater. Currently, linking the size of beneficiary payments to ACO savings would tend to make the least efficient providers more attractive to beneficiaries since those providers are generating the most savings and would offer greater benefits. Therefore, we agree with not establishing that link for now but believe it is worth revisiting once more convergence in spending occurs. *At that point, mechanisms to share savings with beneficiaries also could be used to mitigate incentives for ACOs to select low-risk patients by stipulating that higher-risk patients receive more of the savings than lower-risk patients.* ACOs engaging in favorable risk selection to generate savings would then become more attractive to the patients they may be trying to avoid. We understand that the Bipartisan Budget Act currently prohibits unequal payments to different beneficiaries. We are suggesting this potential policy direction for future consideration.

**Recommendation 7:** *Eliminate cap on downward benchmark adjustments that result from decreases in risk scores to mitigate what would otherwise be strong incentives to engage in favorable risk selection. Use ACO CAHPS data to help distinguish between changes in coding intensity and changes in population health status. Should benchmarks be converged to a regional average in the future (not in the near term per Recommendation 4), implement additional payments or benchmark adjustments so that ACOs are not penalized for serving higher-risk patients.*

## **UPDATING BENCHMARKS**

A concern raised about the current methodology for updating benchmarks is that an ACO's own attributed patients are included in the determination of regional spending trends, causing updates to be based—sometimes in no small part—on an ACO's performance. ACOs that serve a significant proportion of the population in their service area are penalized with smaller benchmark updates when they lower spending, thereby weakening incentives for ACOs to ever lower spending (see discussion of this with respect to rebasing above). Many stakeholders have advocated for removing an ACO's population from the determination of its regional trend. In the proposed rule, CMS noted 3 problems with that approach. The first is that removing the ACO's population will leave smaller sample sizes

for regional trend assessment. But small sample sizes can be handled easily with hierarchical statistical modeling approaches. The second is that ACO populations may differ markedly from non-ACO populations and thus have distinct expected spending trends. But research to date has found that ACO and non-ACO populations are largely similar, on average. Moreover, this concern is best addressed with rigorous risk adjustment of benchmark updates for an ACO's population risk, as risk adjustment will not diminish incentives to save. The third is that removing an ACO's population will strengthen incentives for ACOs to select patients with favorable risks. We believe this, too, is most appropriately addressed with risk adjustment. No matter the updating methodology, ACOs will have incentives to select patients that are lower-risk in ways that are not accounted for by the risk adjustment methodology. As we discuss above, we believe the risk adjustment methodology can be improved and additional steps can be taken to minimize incentives for favorable risk selection.

Therefore, we recommend removing an ACO's attributed population from update determinations. We generally agree with the proposed blending of concurrent national and regional growth rates to update benchmarks, with greater weight given to national growth rates for ACOs in areas where the FFS population served by other providers is limited. But rather than use an arbitrary weighting scheme, we recommend the use of standard statistical approaches to estimation that account for random error from finite sample sizes. Such methods are used, for example, to estimate excess readmission rates for the Hospital Readmissions Reduction Program. Specifically, at the end of each performance year, CMS analysts could use claims data from the performance year and baseline year 3 (BY3) to fit a simple multilevel model of total Medicare spending per beneficiary as a function of county-level random intercepts and (time) slopes, *excluding beneficiaries attributed to the ACO of interest* (the model would be estimated separately to determine each ACO's update). With empirical Bayes estimation, the model will produce estimates of county-level spending trends that are shrunken toward the national trend when a county's sample is small (whether because the county is small or because a large proportion of its population is served by the ACO) and closer to the unshrunk county trend when the population in the county not assigned to the ACO is large. For each ACO, a weighted average of the estimated (shrunk) county trends can then be taken (reflecting the geographic distribution of the ACO's attributed patients) to arrive at the benchmark update for that ACO.

This approach would preserve incentives for ACOs to lower spending by removing their assigned population from update determinations while incorporating regional trends to the extent they can be reliably estimated and flexibly giving more weight to the national trend for ACOs that serve higher proportions of their service area. We believe it is important initially to incorporate regional trend information to the extent possible because it is important initially to align benchmarks with ACOs' counterfactual spending to engender trust that ACOs will be rewarded for improvement. After ACOs have had a few years to lower spending, benchmarks may later diverge from counterfactual spending without reactionary participation losses, as long as the divergence does not depend on the ACO's past performance, as discussed above. This method achieves the goal of incorporating regional trends to the extent it can be reliably supported by the data.

By including the rest of the ACO population in update determinations (all but those assigned to the ACO of interest), *this approach to benchmark updating will still apply downward pressure on ACO benchmarks at a rate that follows, but is less than, the rate of slowing in spending caused by the program. Thus, rebasing is not necessary to exert this downward pressure.* Including the rest of the ACO population in update determination will also introduce an element of competition in markets with multiple ACOs, as ACOs will be rewarded more if they lower spending more than their competitors. This dynamic may become undesirably prominent in large service areas dominated by multiple ACOs. In such markets, the weight given to the national trend will be limited because the market is large (allowing stable trend estimation), so that the update for an ACO will be based almost entirely on the performance of its neighboring ACO competitors. Thus, if all ACOs lower spending significantly and equally, none would be rewarded for doing so. Such cases are not currently common. A more common scenario is the single ACO dominating a small market. For those cases, our recommended methodology would rely on national trends for updates. But CMS could monitor for cases of high ACO penetration in large markets with multiple ACOs and use national trends for those markets, too.

More generally, should the MSSP expand to the point that the national non-ACO population becomes small, we would recommend choosing a desirable rate at which benchmarks should grow and apply that rate to all ACOs. The methodology we describe above is best suited to an expansion phase of the program.

***Recommendation 8:*** *Remove an ACO's attributed population from benchmark update determinations to preserve incentives for ACOs to lower spending. Use multilevel statistical modeling approaches to estimate ACO-specific update factors that reflect regional and national trends, with greater weight placed on national trends for ACO's that serve higher proportions of patients in their service area. After a phase of program expansion, switch to using a desirable preset rate of benchmark growth applied to all ACOs.*

Thank you for your consideration of our comments. If you should wish to discuss any of our analyses or recommendations further, you can reach us at the email address below.



J. Michael McWilliams, MD, PhD  
Warren Alpert Foundation Professor of Health Care Policy, Harvard Medical School  
[mcwilliams@hcp.med.harvard.edu](mailto:mcwilliams@hcp.med.harvard.edu)



Michael E. Chernew, PhD  
Leonard D. Schaeffer Professor of Health Care Policy, Harvard Medical School



Bruce E. Landon, MD, MBA  
Professor of Health Care Policy and Medicine, Harvard Medical School