By David M. Eddy and Roshan Shah

DOI: 10.1377/hlthaff.2012.0385
HEALTH AFFAIRS 31,
NO. 11 (2012): 2554–2562

A Simulation Shows Limited Savings From Meeting Quality Targets Under The Medicare Shared Savings Program

David M. Eddy (eddyaspen@yahoo.com) is founder and chief medical officer emeritus at Archimedes, in San Francisco, California.

Roshan Shah is an associate scientist at Archimedes.

ABSTRACT **The Medicare Shared Savings Program, created under the Affordable Care Act, will reward participating accountable care organizations that succeed in lowering health care costs while improving performance. Depending on how the organizations perform on several quality measures, they will "share savings" in Medicare Part A and B payments—that is, they will receive bonus payments for lowering costs. We used a simulation model to analyze the effects of the Shared Savings Program quality measures and performance targets on Medicare costs in a simulated population of patients ages 65–75 with type 2 diabetes. We found that a ten-percentage-point improvement in performance on diabetes quality measures would reduce Medicare costs only by up to about 1 percent. After the costs of performance improvement, such as additional tests or visits, are accounted for, the savings would decrease or become cost increases. To achieve greater savings, accountable care organizations will have to lower costs by other means, such as through improved use of information technology and care coordination.**

An important initiative under the Affordable Care Act is the Centers for Medicare and Medicaid Services' (CMS's) Shared Savings Program, which provides financial incentives for provider organizations that meet specified standards for quality performance.[1] The program represents a key step toward correcting one of the major problems in the US health care system: a misalignment of incentives that rewards physicians and hospitals for providing the maximum amount of health care services, instead of prioritizing care to maximize value and seeking ways to reduce costs without reducing value. Although nominally developed for Medicare, the Shared Savings Program is also being considered as a model for curbing health care spending for all payers and age groups.[2,3]

To participate in the program, providers must form an accountable care organization, which is a group of providers that is collectively held accountable for the overall cost and quality of care for a defined patient population—in this case, beneficiaries enrolled in Medicare Parts A and B. Depending on the organization's performance on a set of specified measures, providers will share in any savings in Parts A and B payments that may result from preventing adverse events and their associated costs (such as the cost of hospital admissions) or finding more efficient ways to deliver care (such as coordinating visits to manage chronic conditions).

CMS has issued rules that specify how to become designated as an accountable care organization, measure performance, calculate changes in costs, and calculate shared savings.[4] The cost of the drugs that beneficiaries receive under the Medicare Part D prescription drug benefit does not count toward the accountable care organization's cost of caring for the beneficiaries, for the purposes of calculating cost savings.

The rules for becoming an accountable care

organization raise some important questions. To what extent will performance improvements generate meaningful cost savings for such organizations? More specifically, given the program rules[4] and realistic improvements in performance, how many adverse events such as strokes will be prevented, and what savings will this prevention produce in Medicare Parts A and B payments? To what extent will those savings be offset by the costs of the interventions required to achieve performance improvement, such as visits to physicians' offices, tests, and drugs? What will the effects of each of the performance measures be on clinical and cost outcomes? Are the measures equally important—and, if not, how should they be prioritized?

The answers will differ for accountable care organizations depending on their patient populations, current performance levels, and cost structures. However, some general conclusions can be drawn.

For this study we used a tool called the Archimedes model to simulate what would occur if the Shared Savings Program were implemented in an "average" accountable care organization—one that has a population, care processes, and costs representative of Medicare beneficiaries and providers. The Archimedes model simulates the outcomes of individual people with regard to variables specified in the performance measures, such as blood pressure, cholesterol, and glucose levels. It calculates expected health and cost outcomes based on current levels of care, and it calculates the effects of changes in care on these outcomes.

We calculated the effects of improving performance on the accountable care organization measures that are most directly related to quality and cost, using a subpopulation that can be expected to derive the largest benefit: patients with type 2 diabetes. This population is the largest at-risk population covered by the measures, has the highest risk of the widest range of adverse events, and is the target for the largest number of performance measures. Thus, it provides a case study of an area in which the Shared Savings Program can be expected to have the largest impact.

Rules For The Shared Savings Program

The performance measures specified by CMS fall into four domains: patient-caregiver experience; care coordination and patient safety; preventive health; and at-risk populations. Seven of the twelve measures in the at-risk domain are targeted at diabetes care. The measures, whose treatment goals are listed in parentheses, are as follows: hemoglobin A1c (HbA1c) (< 8 percent), low-density lipoprotein (LDL) cholesterol (< 100 mg/dL), blood pressure (< 140/90 mmHg), tobacco (nonuse), aspirin (use), a composite of the first five measures (all five preceding goals), and HbA1c poor control (< 9 percent).[4,5]

Although all seven measures will be reported to CMS, only the composite and HbA1c poor control measures will be scored by CMS for the purposes of calculating shared savings. The two measures will be weighted equally in these calculations.

Each of the two measures will be assigned a score that gives minimum credit if only a "threshold" level of performance is achieved and that gives maximum credit if a "benchmark" level of performance is achieved. CMS will define both the threshold and benchmark levels. These scores will determine the proportion of savings that the accountable care organization will receive.

In the first three years, organizations will be able to choose between a "one-sided" option, in which they can receive up to 50 percent of savings if costs decrease but will not be at risk if costs increase, or a "two-sided" option, in which they can receive up to 60 percent of any savings but could be penalized if costs increase. After three years all organizations will use the two-sided option. In the one-sided option, the savings in any particular year must exceed 2 percent before shared savings kick in, to ensure that the decreases in cost are not the result of random fluctuations.

CMS has not yet specified the thresholds or benchmarks. More information on CMS's methods is included in the online Appendix.[6]

Study Data And Methods

THE SIMULATION To analyze the possible effects of this program on health and cost outcomes, we created a simulated Medicare patient population, with simulated providers who follow protocols specifying how this population should be treated. We ran simulations under eight different scenarios, as follows: a "control," in which current levels of care continue; and a set of performance improvement scenarios, in which performance on each of the seven performance measures, one by one, was improved to a specified level.

We used the Archimedes model[7,8] to conduct the simulations, and we used the ARCHeS web interface tool[9] to specify the population and clinical scenarios. The Archimedes model uses equations to represent physiological pathways and the occurrence and progression of diseases.

It includes the occurrence of symptoms; patients' encounters with the health care system, such as visits to providers' offices and hospital admissions; providers; care processes, such as guidelines and protocols; providers' and patients' behaviors; interventions, such as prevention programs, tests, procedures, and treatments; health outcomes; costs; and quality of life. It also includes tools for creating simulated populations that match real populations.

The structure of the model, variables, methods for deriving equations, and data sources are described elsewhere.[10] We validate the model by simulating epidemiological studies, clinical trials, and other empirical studies.[11] The Archimedes model has been used to conduct a variety of analyses spanning several diseases,[12] including many analyses of performance measures and guidelines relating to diabetes and cardiometabolic risk—that is, cardiovascular disease and diabetes.

For this analysis we created a simulated Medicare population ages 65–75—the upper age limit captured by the accountable care organization performance measures—using methods described elsewhere[10,11] and data from the National Health and Nutrition Evaluation Survey.[13] Current performance levels were obtained from the same survey, and care processes were based on US national guidelines,[10] calibrated to match data in US national surveys[14–16] using methods described elsewhere.[10] The effects of treatments were based on clinical trial results.[10] Costs for all aspects of care except the interventions targeted by the performance measures were based on Medicare costs.[17,18]

For each quality measure, we calculated the effects of a ten-percentage-point improvement in performance. We assumed that the improvements in performance would come from a representative sample of patients who were candidates for performance improvement, such as a representative sample of those with HbA1c greater than 8 percent. We calculated the effects of each measure on myocardial infarctions, strokes, microvascular complications of diabetes (retinopathy, neuropathy, and nephropathy), life-years, and quality-adjusted life-years. For the composite measure, we assumed that the ten-percentage-point improvement would be accomplished by successfully treating randomly selected patients who currently do not meet the treatment goal of one or more of the component measures.

We then calculated the costs of the additional visits, tests, and drugs needed to bring these patients to the desired goal through treatment. We assumed that patients would need between one and three additional visits per year, depending on the measure, and would receive the tests recommended in national guidelines.

We used National Health and Nutrition Evaluation Survey data[13] on currently treated patients to determine the proportion receiving various drugs, and we assumed that newly treated patients would be treated similarly. Costs of visits, tests, and drugs were based on Medicare data[19,20] or other data[21] when Medicare data were not available. Specific cost assumptions are listed in the online Appendix.[6]

We calculated results over a five-year period to learn the effects in the first few years of the Shared Savings Program. Because savings are calculated annually, we report the results on an annualized basis.

STUDY LIMITATIONS Our analysis has several limitations. We calculated outcomes only for diabetes and its complications; the additional benefits to other conditions, such as the effects of smoking cessation on lung disease, were not included.

In addition, to determine the benefits of the measures as they are designed, we calculated the effects of treating biomarkers (HbA1c, LDL, and blood pressure) to exactly the goals specified by the measures. In reality, when patients reach a goal through treatment, they usually improve beyond the exact goal specified in the measure.

Conversely, we assumed that all patients could reach the goal with treatment. This may be overly optimistic, especially for organizations that already have aggressive performance improvement programs, and especially for the smoking cessation measure. Taken together, these two assumptions tend to balance each other out.

Additional limitations are discussed in the Appendix.[6]

Our results are subject to several sources of variation and uncertainty. The results for particular accountable care organizations will vary depending on factors such as the distribution of risk factors in their patient populations, current levels of performance, actual increases in performance that they will be able to achieve, and actual costs of the interventions needed to improve performance. An organization's share of any savings will be affected by the threshold and benchmark levels of performance that CMS eventually sets. The range of uncertainty around the outcomes will be affected by the size of the organization's population—with random variations affecting smaller organizations more than larger ones—and the sizes of the subpopulations targeted by particular measures.

Study Results

Exhibit 1 shows the characteristics of the pa-

EXHIBIT 1

Characteristics Of Medicare Patients With Diabetes In Archimedes Simulation

Patients with diabetes	Percent of total Medicare diabetes population	Average SBP	Average LDL	Average HbA1c	Smoke (%)	Percent taking treatment: Aspirin	Antihypertensive medication	Dyslipidemia medication	Oral diabetes medication	Insulin
All	100	130.4	107.9	7.3	7.8	52.3	65.1	56.1	70.9	22.4
Not meeting CMS performance measure										
HbA1c > 8%	22	132.0	109.9	9.5	7.4	55.9	66.4	57.4	64.6	0.4
LDL >100	34	131.3	145.5	7.3	9.3	52.7	63.1	18.7	68.6	22.0
Not on aspirin	48	130.5	108.0	7.2	7.1	0.0	65.1	55.6	71.7	21.8
SBP >140	22	153.6	109.6	7.3	8.3	50.4	57.7	56.4	66.4	27.8
Smoke	8	130.1	111.6	7.2	100.0	57.0	65.9	53.4	62.7	23.7
Composite	62	134.7	120.1	7.7	12.6	53.1	63.5	44.1	67.4	20.2
HbA1c poor control (> 9%)	12	132.1	107.7	10.4	8.0	58.4	70.0	60.1	55.0	0.3

SOURCE Archimedes model, based on data from the National Health and Nutrition Evaluation survey (see Note 13 in text). **NOTES** The performance measures of the Centers for Medicare and Medicaid Services (CMS) are explained in the text. Composite measure does not include HbA1c poor control (hemoglobin A1c >9 percent). HbA1c is measured as a percentage. Systolic blood pressure (SBP) is measured in mmHg. Low-density lipoprotein (LDL) cholesterol is measured in mg/dL.

tients who are candidates for performance improvement on each measure—that is, those who have not met the measure's goal. It provides information about which groups to focus performance improvement activities on, as well as the feasibility of improving performance. There are relatively few patients in the study population with HbA1c greater than 9 percent (12 percent of the diabetes population), compared to the larger number of patients who are candidates for the composite measure (62 percent).

Exhibit 1 also indicates the size of the biomarker gap that must be closed. For example, of patients with HbA1c greater than 8 percent, the average HbA1c is 9.5 percent, which indicates that an average reduction of 1.5 percentage points will be required. For patients with systolic blood pressure greater than 140 mmHg, the average reduction will need to be 13 mmHg.

Also notable are the proportions of patients not yet taking treatment. Of patients with HbA1c greater than 8 percent or greater than 9 percent, fewer than 0.5 percent are on insulin, compared to 22.4 percent for the overall Medicare diabetes population. Of patients with LDL greater than 100 mg/dL, only 18.7 percent are on dyslipidemia medication, compared to 56.1 percent in the overall diabetes population. Similarly, 42.3 percent of patients with systolic blood pressure greater than 140 mmHg are not on antihypertensive medication.

Exhibit 2 shows current performance levels and the effects on health outcomes of a ten-percentage-point increase in performance for each of the measures. Because 92.2 percent of the patients in this population are already nonsmokers, a ten-percentage-point improvement would require that no patient smoked. Performance on the HbA1c poor control measure (> 9 percent) is also already very high, with 88.3 percent having HbA1c less than 9 percent. Thus, a ten-percentage-point improvement on this measure would require successful treatment of almost everyone in that group.

As expected, improving performance on the composite measure has the largest effect—for example, preventing 4.13 percent of myocardial infarctions in the Medicare diabetes population. Smoking cessation reduces both myocardial infarctions and strokes by more than 3 percent each. The effects of improving performance on the other measures are smaller.

Exhibit 3 shows the effects of a ten-percentage-point performance improvement on costs. Across all measures, the savings that can be expected from preventing adverse events are all 1.22 percent or lower. This is below the 2 percent threshold that CMS has set for sharing savings with providers. As with health outcomes, the cost effects are relatively small because each performance measure targets only a subset of the diabetes population, and we have assumed an increase in performance of only ten percentage points.

When the costs of the visits and tests (covered under Part B) required to improve performance are factored in to produce the total net effect on costs, the savings to Parts A and B are reduced or

EXHIBIT 2

Impact On Annual Health Outcomes If Affordable Care Organizations Improved Performance By Ten Percentage Points

	Performance		Change in health outcomes				
Patients with diabetes	**Current**	**Improved**	**Myocardial infarctions**	**Strokes**	**Microvascular complications**	**Life-years**	**QALYs**
All	—[a]	—[a]	1,352	774	9,456	93,408	75,386
Not meeting CMS performance measure							
HbA1c > 8%							
Change	78.4%	88.4%	−32	−3	−128	36	38
Percent change	—[a]	—[a]	−2.37	−0.39	−1.35	0.04	0.05
LDL >100							
Change	53.0%	63.0%	−28	0	0	12	11
Percent change	—[a]	—[a]	−2.08	0.00	0.00	0.01	0.01
Not on aspirin							
Change	52.3%	62.3%	−20	−3	5	19	18
Percent change	—[a]	—[a]	−1.46	−0.34	0.06	0.02	0.02
BP >140/90							
Change	72.5%	82.5%	−18	−5	−219	47	53
Percent change	—[a]	—[a]	−1.35	−0.59	−2.32	0.05	0.07
Smoke							
Change	92.2%	100.0%	−50	−25	−94	43	41
Percent change	—[a]	—[a]	−3.72	−3.26	−1.00	0.05	0.05
Composite							
Change	38.0%	48.0%	−56	−11	−149	50	53
Percent change	—[a]	—[a]	−4.13	−1.45	−1.58	0.05	0.07
HbA1c poor control (> 9%)							
Change	88.3%	98.3%	−32	0	−113	49	47
Percent change	—[a]	—[a]	−2.38	0.00	−1.19	0.05	0.06

SOURCE Archimedes model, based on data from the National Health and Nutrition Evaluation survey (see Note 13 in text). **NOTES** Outcomes shown for 100,000 Medicare patients with diabetes followed for five years, with results annualized. The three main microvascular complications are retinopathy, neuropathy, and nephropathy. QALY is quality-adjusted life-year. [a]Not applicable.

reversed (Exhibit 3). Only the composite, blood pressure, aspirin, and smoking measures show a net savings to Parts A and B after these activities are taken into account, with the savings ranging from 0.02 percent (aspirin) to about 1 percent (smoking). Drugs taken by these patients on an outpatient basis are not covered by Part A or Part B, but their costs may be paid by Medicare under the Part D benefit; by accountable care organizations; or by other insurers under Medicare Advantage plans, perhaps with copayments by patients.

Exhibit 3 also shows the effects of including all types of these costs, including drugs. Only the blood pressure and smoking cessation measures save money when all costs are considered, with ten-percentage-point improvements on these measures leading to reductions in overall costs of 0.11 percent and 0.78 percent, respectively.

We created a spreadsheet (available from the authors on request) for exploring the effects of different assumptions about costs and performance, such as the expected impact of performance improvements greater or less than ten percentage points. Results for different populations other than the beneficiaries covered in this study also can be calculated using the ARCHeS web interface.[9] However, under any reasonable set of assumptions, the annual savings in Parts A and B for the study population are likely to fall below CMS's 2 percent minimum threshold for sharing savings under the one-sided option.

Discussion

The Shared Savings Program is a landmark achievement, fostering the creation of organizations that are accountable for both the quality and cost of the care they provide and giving them an incentive that links the two: the greater the improvement in their quality, the greater their share of any savings. A critical component of the program is the specification of performance measures and the direct linking of improvements in performance to shared savings for accountable care organizations.

A fundamental question is the extent to which feasible improvements in performance by themselves can be expected to produce substantial savings. If they can, accountable care organiza-

EXHIBIT 3

Impact On Annual Costs (In Millions Of Dollars) If Affordable Care Organizations Improved Performance By Ten Percentage Points

Patients with diabetes	Effect on Parts A and B costs from preventing CVD events	Cost of interventions to improve performance		Net effect on costs		
		Visits and tests	Drugs	Parts A and B	Drugs	Total
All	$1,033.1	—[a]	—[a]	$1,033.1	$228.3	$1,261.4
Not meeting CMS performance measure						
HbA1c > 8%						
Change	−$2.127	$3.726	$9.622	$1.599	$9.622	$11.221
Percent change	−0.21	—[a]	—[a]	0.15	4.22	0.89
LDL >100						
Change	−$0.911	$2.383	$8.126	$1.472	$8.126	$9.597
Percent change	−0.09	—[a]	—[a]	0.14	3.56	0.76
Not on aspirin						
Change	−$1.115	$0.921	$0.572	−$0.194	$0.572	$0.377
Percent change	−0.11	—[a]	—[a]	−0.02	0.25	0.03
BP >140/90						
Change	−$8.781	$2.271	$5.180	−$6.510	$5.180	−$1.330
Percent change	−0.85	—[a]	—[a]	−0.63	2.27	−0.11
Smoke						
Change	−$12.621	$2.350	$0.470	−$10.271	$0.470	−$9.801
Percent change	−1.22	—[a]	—[a]	−0.99	0.21	−0.78
Composite						
Change	−$7.742	$5.574	$12.188	−$2.168	$12.188	$10.020
Percent change	−0.75	—[a]	—[a]	−0.21	5.34	0.79
HbA1c poor pontrol (> 9%)						
Change	−$2.086	$2.791	$9.609	$0.705	$9.609	$10.314
Percent change	−0.20	—[a]	—[a]	0.07	4.21	0.82

SOURCE Archimedes model, based on data from the National Health and Nutrition Evaluation survey (see Note 13 in text). **NOTES** Dollar figures represent millions of dollars, and individual items might not sum to the total because of rounding. Outcomes are shown for 100,000 Medicare patients with diabetes followed for five years, with results annualized. Parts A and B costs include the cost of both the visits and the tests needed to improve performance on the measures and all other Parts A and B costs—in other words, the costs that are eligible to be "shared" between Medicare and the accountable care organization. The performance measures of the Centers for Medicare and Medicaid Services (CMS) are explained in the text. Composite measure does not include HbA1c poor control (hemoglobin A1c >9 percent). HbA1c is measured as a percentage. Blood pressure (BP) is measured in mmHg. Low-density lipoprotein cholesterol (LDL) is measured in mg/dL. CVD is cardiovascular disease. [a]Not applicable.

tions will have strong incentives to improve performance and capture the resulting savings. If not, then performance will still be important in determining an organization's share of any savings, but the actual savings will have to come from other activities, as discussed below.

Of the performance measures, only the domains of preventive health and at-risk populations, which stimulate the use of evidence-based treatments, will have direct effects on health outcomes and their associated costs. The other two domains—patient-caregiver experience, and care coordination and patient safety—will improve the quality of care, but any effects they may have on health outcomes will be indirect and impossible to predict.

To understand the potential effects on savings of improving performance, we examined performance measures for a population most likely to demonstrate such effects: patients with diabetes. Because of the size of this population, the high risk of adverse events, and large number of effective treatments, the hoped-for effects of the Shared Savings Program would be most likely to appear in and be greatest in this group of patients.

The measures chosen by CMS for determining the quality of care for diabetes patients all promote treatments that have important effects on the measures' respective target populations. From the perspective of the total population of patients with diabetes, however, the measures' effects are muted by the fact that each measure addresses only a subset of the population, and that performance will not start at 0 percent or jump to 100 percent. For our baseline analysis, we chose a ten-percentage-point improvement in performance on each measure. This is aggressive but feasible because for every measure there is a pool of currently untreated patients (Exhibit 1).

POTENTIAL SAVINGS Our analysis indicated that under our baseline assumptions, a ten-percentage-point improvement in performance will prevent up to 4.1 percent of adverse events—such as strokes; myocardial infarctions; and microvascular complications, including retinopathy, neuropathy, and nephropathy—which in turn will create savings in Parts A and B of up to 1.22 percent, depending on the measure. However, the savings are diminished or become cost increases when the cost of the visits and tests needed to improve performance is included.

The net effect is that the savings in Parts A and B, if any, are very likely to fall below the 2 percent limit that CMS has set for sharing savings in the one-sided option. Furthermore, accountable care organizations would receive only 50 percent or 60 percent of their achieved savings, depending on whether they were participating in the one- or the two-sided option.

In addition, all of the interventions require prescribing drugs to beneficiaries. Even if drug costs are not included in the calculation of Part A or Part B savings, they will be paid by different parties depending on the individual financial arrangements involved. These costs can be large compared to the costs saved by preventing downstream events (Exhibit 3). Accountable care organizations will also face start-up costs and first-year operating costs, which have been estimated by CMS to be about $1.7 million per organization.[22] The organizations' savings or increases in costs will depend heavily on the threshold or benchmark levels of performance that CMS will set.

Given that the savings that can be expected from preventing downstream events will be small or nonexistent, accountable care organizations will need to cut costs in other ways. As a result, the main motivation for improving performance on the Shared Savings Program performance measures will not be to create immediate savings from those efforts, but rather to gain a larger share of whatever savings can be found elsewhere.

USING ONLY TWO MEASURES TO CALCULATE SAVINGS The two diabetes measures that CMS will score for the purposes of calculating the share of savings—the HbA1c poor control ($>$ 9 percent) measure and the composite measure—are currently weighted equally in CMS's calculation of shared savings. However, the measures vary widely in their respective importance to clinical outcomes. Improvements on the composite measure have a far larger effect on health and economic outcomes than similarly sized improvements in the HbA1c poor control measure.

The threshold and benchmark levels of performance that CMS sets will also affect the relative effects of the two measures. For instance, if achieving a ten-percentage-point improvement in performance provides "full credit" in the scoring system, then the composite measure will be about twice as important as the HbA1c poor control measure in terms of preventing myocardial infarctions and strokes. In terms of Parts A and B costs, a ten-percentage-point improvement in the composite measure would reduce costs about 0.21 percent, whereas a similar improvement in the HbA1c poor control measure would increase costs about 0.07 percent (Exhibit 3).

If instead full credit requires achieving a 90 percent level of performance—the default specified by CMS if no better data are available[4]—then the composite measure would be vastly more important in reducing myocardial infarctions and strokes than the HbA1c poor control measure (281 events prevented, versus 6). Requiring this level of improvement would create an even wider gap in the Parts A and B costs of the two measures, with an increase in cost of 0.01 percent for the HbA1c poor control measure versus a decrease in cost of 1.09 percent for the composite measure.

Relying on the composite measure also carries the risk that an accountable care organization could put resources into one or more of its component measures and succeed in reaching those goals but not succeed on the other components and ultimately get no credit for its efforts. This could pose a disincentive to work on the component measures, some of which—for example, controlling hypertension and smoking—represent powerful ways to improve outcomes and reduce costs. Consideration should be given to disaggregating the component measure and returning to the original proposal[22] of including all measures in the score individually. Different weights could then be assigned to different measures based on their potential effects on health and economic outcomes.

Conclusions

CMS has projected that in the first three years of the Shared Savings Program, accountable care organizations will receive estimated median shared savings of $800 million spread among all participating organizations (tenth and ninetieth percentile uncertainty levels: $560 and $1,130 million, respectively).[22] Given the agency's assumption that 75 to 150 organizations will participate, this corresponds to shared savings payments to each organization of about $2.5 million per year. Our analysis indicates that the savings needed to generate these payments will have to come from activities other than im-

provements in the clinical quality measures.

CMS anticipates that accountable care organizations will "improve upon information technology systems, focus on evidence-based medicine, improve care coordination and quality and generally refine all processes of caring for their patients and community."[22(p19639)] The results of the Physician Group Practice Demonstration, considered the prototype for the accountable care organization program,[23] support our contention that the savings will have to come from nonmeasured activities. In that program, only six of ten primary care practices achieved any shared savings, and they attributed their success to infrastructural and environmental factors such as organizational structure; investments in care management programs and redesigned care processes; more intensive diagnostic coding; changes in market conditions; and previously favorable cost trends.

One practice, Marshfield Clinic of Wisconsin, was responsible for about 50 percent of the savings across all ten practices. The director of Marshfield's activities ascribed the clinic's success to an aggressive acceleration of "multiple initiatives…including health information technology (point-of-care reminders, being completely chartless), care management programs, [and] education and feedback to providers regarding populations of patients with a given condition."[23(p199)] It is these types of general activities, rather than improvements in performance on the clinical measures alone, that will help generate the desired savings.

The performance levels of accountable care organizations will be important primarily for determining the share of those other savings that the organizations will be able to keep. If they are already high performers compared to the threshold and benchmark levels that CMS sets, organizations may choose to focus less on performance measure improvements and more on finding efficiencies elsewhere in their programs.

When estimating the overall effects on their programs, accountable care organizations and CMS should analyze the cost of improving performance, including drug costs even though they are not covered by Part A or Part B. As noted above, CMS should consider disaggregating the component measure and applying different weights to different measures based on their potential effects on health and economic outcomes.

CMS will need to be very careful in setting the threshold and benchmark levels because they could profoundly affect the outcomes of the Shared Savings Program. Over time, CMS should consider an incentive based on the overall effect on health outcomes, not just on individual performance measures. ■

The authors thank the Robert Wood Johnson Foundation for funding the development of the ARCHeS interface, which makes the Archimedes model accessible online. [*Published online October 3, 2012.*]

NOTES

1 CMS.gov. Shared Savings Program [Internet]. Baltimore (MD): Centers for Medicare and Medicaid Services; [last modified 2012 May 29; cited 2012 Sep 18]. Available from: https://www.cms.gov/sharedsavingsprogram/

2 Accountable Care Organization Learning Network. ACO pilot sites [Internet]. Washington (DC): Brookings Institution; c2012 [cited 2012 Sep 18]. Available from: http://www.acolearningnetwork.org/what-we-do/aco-pilot-sites

3 Center for Medicare and Medicaid Innovation. Pioneer ACO Model [Internet]. Baltimore (MD): Centers for Medicare and Medicaid Services; [cited 2012 Sep 18]. Available from: http://innovations.cms.gov/areas-of-focus/seamless-and-coordinated-care-models/pioneer-aco/

4 Centers for Medicare and Medicaid Services. Medicare program: Medicare Shared Savings Program: accountable care organizations: final rule. Fed Regist. 2011;76(212):67802–990.

5 RTI International, Telligen. Accountable care organization 2012 program analysis, quality performance standards narrative measure specifications, final report. Research Triangle Park (NC): RTI International; 2011 Dec 12. (RTI Project No. 0213195.000.004).

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

7 Eddy DM, Schlessinger L. Archimedes: a trial-validated model of diabetes. Diabetes Care. 2003;26(11):3093–101.

8 Eddy DM, Schlessinger L. Validation of the Archimedes diabetes model. Diabetes Care. 2003;26(11):3102–110.

9 Archimedes. Online healthcare forecasts: ARCHeS [Internet]. San Francisco (CA): Archimedes; [cited 2012 Sep 19]. Available from: http://archimedesmodel.com/arches-services

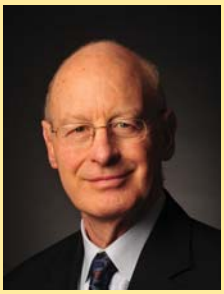
10 Eddy DM, Cohen M, editors. Description of the Archimedes model, simulator 2.3 [Internet]. San Francisco (CA): Archimedes; 2011 Dec [cited 29 Apr 2012]. (Archimedes Report). Available from: http://www.archimedesmodel.com/tech-reports

11 Eddy DM, Cohen M, Shum K, editors. Validation methodology results, Archimedes Simulator 2.3 [Internet]. San Francisco (CA): Archimedes; 2011 Dec [cited 29 Apr 2012]. (Archimedes Report). Available from: http://www.archimedesmodel.com/tech-reports

12 For a list of publications, see Archimedes. Facts and figures: resource center: scientific publications [In-

ternet]. San Francisco (CA): Archimedes; [cited 2012 Sep 19]. Available from: http://archimedesmodel.com/publications
13 Centers for Disease Control and Prevention. National Health and Nutrition Examination Survey [Internet]. Atlanta (GA): CDC; [last updated 2012 Sep 26; cited 2012 Sep 27]. Available from: http://www.cdc.gov/nchs/nhanes.htm
14 Centers for Disease Control and Prevention. Ambulatory health care data [Internet]. Atlanta (GA): CDC; [last updated 2012 Sep 18; cited 2012 Sep 19]. Available from: http://www.cdc.gov/nchs/ahcd.htm
15 Centers for Disease Control and Prevention. National Hospital Discharge Survey [Internet]. Atlanta (GA): CDC; [last updated 2012 Sep 27; cited 2012 Sep 27]. Available from: http://www.cdc.gov/nchs/nhds.htm
16 Centers for Disease Control and Prevention. Compressed Mortality File [Internet]. Atlanta (GA): CDC; [last updated 2012 Jul 12; cited 2012 Sep 19]. Available from: http://www.cdc.gov/nchs/data_access/cmf.htm
17 CMS.gov. Medicare Current Beneficiary Survey (MCBS) [Internet]. Baltimore (MD): Centers for Medicare and Medicaid Services; [last modified 2012 Jul 12; cited 2012 Sep 19]. Available from: https://www.cms.gov/MCBS
18 CMS.gov. Limited Data Set (LDS) files [Internet]. Baltimore (MD): Centers for Medicare and Medicaid Services; [last modified 2012 Jun 21; cited 2012 Sep 19]. Available from: https://www.cms.gov/LimitedDataSets
19 CMS.gov. Standard analytical files—LDS [Internet]. Baltimore (MD): Centers for Medicare and Medicaid Services; [last modified 2012 Mar 17; cited 2012 Sep 19]. Available from: https://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/LimitedDataSets/StandardAnalyticalFiles.html
20 CMS.gov. CMS guide to requests for Medicare Part D Prescription Drug Event (PDE) data [Internet]. Baltimore (MD): Centers for Medicare and Medicaid Services; 2008 [cited 2012 Sep 30]. Available from: http://www.cms.gov/Medicare/Prescription-Drug-Coverage/PrescriptionDrugCovGenIn/Downloads/GuidePartDv3-3-17-09-2.pdf
21 Curry SJ, Grothaus LC, McAfee T, Pabiniak C. Use and cost effectiveness of smoking-cessation services under four insurance plans in a health maintenance organization. N Engl J Med. 1998;339(10):673–9.
22 Centers for Medicare and Medicaid Services, Office of Inspector General. Medicare program; Medicare Shared Savings Program: accountable care organizations and Medicare program: waiver designs in connection with the Medicare Shared Savings Program and the innovation center; proposed rule and notice. Fed Regist. 2011;76(67):19528–54.
23 Iglehart JK. Assessing an ACO prototype—Medicare's Physician Group Practice demonstration. N Engl J Med. 2011;364(3):198–200.

ABOUT THE AUTHORS: DAVID M. EDDY & ROSHAN SHAH



David M. Eddy is founder and chief medical officer emeritus at Archimedes.

In this month's *Health Affairs*, David Eddy and Roshan Shah report on the simulation they produced of the impact of meeting quality targets under the Medicare Shared Savings Program. Under that program, participating health care organizations will be rewarded with a portion of the savings if they succeed in lowering the costs of treating Medicare beneficiaries while improving performance. Applying the Archimedes model to a simulated population of diabetes patients, Eddy and Shah found that a ten-percentage-point improvement in meeting quality metrics for this population would reduce Medicare costs by only about 1 percent—and that organizations would need to take other steps to achieve greater cost reduction and meet the 2 percent threshold for shared savings.

Eddy is founder and chief medical officer emeritus at Archimedes, which operates the full-scale simulation model that he and his colleagues developed of human physiology, diseases, behaviors, interventions, and health care systems. Since switching career paths from cardiovascular surgery to applied mathematics, Eddy has focused on applying quantitative reasoning and methods to clinical medicine.

Prior to founding Archimedes, Eddy built models to analyze cancer programs in the United States and abroad, which informed evidence-based guidelines for cancer screening, and worked on Bayesian methods for meta-analysis. His previous positions include senior adviser at Kaiser Permanente, chief scientist at the BlueCross BlueShield Association's Technology Evaluation Center, and professorships at Stanford and Duke Universities. Eddy received a medical degree from the University of Virginia and a doctorate in engineering economic systems from Stanford University.

Roshan Shah is an associate scientist at Archimedes.

Shah is an associate scientist at Archimedes and part of the development team of ARCHeS, a web interface to the Archimedes model, which will enable the public to run clinically realistic virtual trials on any population. Prior to joining the company in 2007, Shah was a research associate at FFA Sciences, a molecular test developer. Shah holds a master's degree in biomedical engineering from the University of Southern California.