Integrated Healthcare Association

Innovation through Collaboration

Issue Brief

No. 22 • November 2016

Charting a Course to Value in Physician Group Payment: Key Pay-for-Performance Design Decisions

Jill M. Yegian, Ph.D, Senior Vice President, Programs and Policy
Lindsay Erickson, M.S.P.H., Director, Value Based P4P
Dolores Yanagihara, M.P.H., Vice President, Performance Measurement

In the late 1990s, when health care quality had barely surfaced on the national policy radar screen, California had no fewer than four overlapping and dissonant public physician organization (PO) quality report cards. The discordant quality reporting confused consumers, while the burden and costs of duplicative reporting requirements frustrated providers.[1] Recognizing an unsustainable status quo, health plans and physician organizations worked with the Integrated Healthcare Association (IHA) to launch a large-scale standardized pay-for-performance (P4P) initiative in 2001, supported by funding from the California Health Care Foundation.

At first aimed solely at quality improvement, IHA's P4P program design now emphasizes value by including utilization (resource use) and cost measures alongside quality measures. IHA's Value Based P4P program has evolved into one of the country's largest alternative payment models by merging quality, patient experience, resource use, and cost measures into a single incentive program across multiple health plans. The program relies on shared savings from more efficient care delivery to reward physician organizations performing well on both quality and cost metrics. Today, participation in IHA's Value Based P4P program includes 10 health plans and more than 200 California physician organizations caring for about 9.6 million Californians enrolled in commercial health maintenance organization (HMO) and point of service (POS) products; over $550 million has been paid out since inception. As purchasers, payers, and providers work to retool the delivery system to provide high-quality, affordable, patient-centered care—or high-value care—examining key design and implementation decisions in the IHA Value Based P4P program can inform similar efforts across the country.

Navigating from Quality to Value

In 2001, IHA, working with California health plans and physician organizations, launched a statewide P4P initiative with the goal of creating compelling incentives to drive improvements in clinical quality and patient experience. While steady quality improvements accrued, dramatic increases in health care costs overshadowed quality gains.

In 2010, IHA P4P stakeholders called for modification of the P4P program to reward performance on quality, patient experience, cost, and utilization measures in an integrated fashion. The Value Based P4P program (VBP4P) was developed, and today includes 10 health plans and more than 200 California physician organizations caring for 9.6 million Californians enrolled in commercial HMO and POS plans—one of the largest alternative payment models in the country. At its core, Value Based P4P is based on shared savings, adjusted for performance on quality measures. The overall VBP4P design is intended to improve quality, decrease health care costs, and reward physician organizations for delivering higher-value care (see page 2 for a brief description of core features of the IHA Value Based P4P program).

This issue brief examines key IHA and stakeholder design decisions made during the transition from a quality-based to value-based P4P program. The intent is to inform other payers, providers, and stakeholders as they navigate a

Based in Oakland, Calif., the nonprofit Integrated Healthcare Association (IHA) convenes diverse stakeholders—including physicians, hospitals and health systems, purchasers, and health plans—committed to high-value, integrated care that improves quality and affordability for patients across California and the nation.

www.iha.org • 500 12th Street, Suite 310 • Oakland, Calif. 94607 • (510) 208-1740
© 2016 Integrated Healthcare Association

course toward value-based purchasing and health care delivery. Unlike mandatory programs that define measures and targets from the top down and penalize physicians for falling short, IHA's voluntary and collaborative approach relies on broad and sustained engagement and buy-in among stakeholders. Given that stakeholder interests are not always aligned within and across organizations, each of the major design elements of the program required careful consideration of tradeoffs and implications. Based on review of key incentive design decisions made in the transition from a quality-specific program to a broader focus on value, five major decisions emerge:

- Should incentives reward performance improvement or attainment?
- Should the program reward performance on utilization (resource use) or total cost of care (TCC)?
- How should benchmarks and targets be set to motivate participants?
- How—and how much—should the program focus on supporting participant use of results for improvement?
- To what extent should the program encourage public use of the results?

A consistent principle woven into the design decisions is the importance—and elusiveness—of simplicity. Every major decision began with clear and strong intent to keep things simple—a laudable goal but challenging in practice. Simpler is easier to understand but lacks nuance. For example, strict application of simple criteria could lead to rewarding POs that are able to achieve a reduction in total cost of care—but from a very high baseline. Each logical decision to adjust the design to ensure broad engagement and credibility among participants adds complexity. In many cases, established approaches and methodologies exist. A critical lesson from Value Based P4P is to identify existing approaches and resist tinkering where possible, carefully considering each decision to ensure that the incremental benefit merits the incremental complexity.

Beyond the five key design decisions outlined, perhaps, the most important element in charting a course toward value in physician payment is creating a program governance structure that is credible to payers, providers, purchasers, policymakers, and the public. Over the last 15 years, the IHA P4P program has built a tried-and-true participant-driven governance structure that stresses balancing individual and shared interests to reach workable compromises.

While labor intensive, the three IHA VBP4P committees—Governance, Technical Measurement, and Technical Payment—provide critical forums for working through myriad details and managing pushback from all quarters. The committees include practicing physicians to ensure that the reality and complexity of incorporating performance measurement in the delivery of care is considered in the design decisions. Other key governance elements include a defined and transparent process that allows all participants to be heard during a public comment period and to pursue recourse through an appeals process. There is nothing easy about reaching common ground on the details of what constitutes "value;" a trusted governance structure is essential so that results are viewed as credible and decisions are perceived as consensus based.

Core Features of IHA Value Based P4P Incentive Design

The VBP4P design is upside risk only; physician organizations (POs) can earn an incentive through shared savings but bear no downside financial risk. To be eligible to earn any share of savings, POs must first meet minimum quality and total cost of care (TCC) standards, known as the Quality Gate and TCC Gate. POs that do not meet these thresholds are ineligible for incentives.

Savings are earned by improvements in resource use, based on five common utilization measures: inpatient bed days, readmissions, emergency department visits, outpatient procedures, and generic prescribing. For example, a reduction in inpatient bed days from 120 to 115 bed days per 1,000 member years reflects improvement of five days for every 1,000 members. The observed improvement is translated into savings based on the cost of a bed day and the size of the physician organization's patient population. Incentive amounts are summed across all measures. Any savings generated are divided between the health plan and the PO. To maximize the incentive earned, POs must improve resource use while simultaneously offering higher quality.

Improving resource use year to year is challenging, particularly for POs that are ahead of the curve. To ensure that POs with efficient resource use are appropriately rewarded, a complementary attainment incentive was added to the design in 2015. This methodology builds in a supplement to the shared-savings calculation for POs that consistently outperform population benchmarks for resource use.

For more information, visit http://www.iha.org/sites/default/files/resources/vbp4p_incentive_design.pdf.

Design Decision No. 1: Improvement vs. Attainment

Incentives that reward performance can focus on improvement or attainment. The initial value-based incentive design,

finalized in 2012 for measurement year (MY) 2013, focused on improvement for two reasons. First, rewards based on improvement motivate broader participation than rewards tied to attainment of specific thresholds, which focus on top performers. With attainment awards, lower-performing physician organizations are likely to disengage if they perceive that the targets are out of reach. Second, an improvement approach allows the program to generate incentive dollars from shared savings—a major advantage from the health plan perspective. A shared-savings approach—an "earned" payment rather than a bonus—was consistent with a guiding program principle of holding down health care costs.

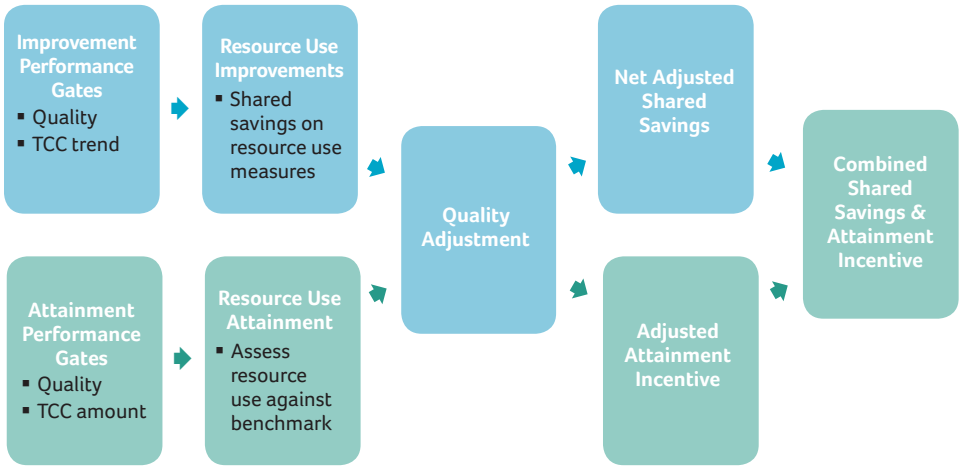
While the initial decision to focus the VBP4P design on improvement (shared savings) was carefully considered and vetted, the first year of implementation and initial payments surfaced concerns that top-performing POs would be ineligible for incentive payments under the new design. Before the transition to the value-based design, over 90 percent of POs received quality-based bonus payments—even if small. The first three health plans that implemented the new design paid bonuses to between 32 percent and 55 percent of POs—a major reduction in the proportion of POs receiving incentives. One reason for the drop was that medical groups and independent practice associations (IPAs) that had already achieved substantial reductions in inpatient bed days and readmissions—key drivers of shared savings under the improvement-oriented incentive design—faced a high bar for continued reductions and were at risk of disengaging. Commitment to the guiding program principle of retaining broad participation led to the incorporation of an attainment incentive in 2015.

The resulting design features two pathways to earn incentive payments: through *improvement incentives* (shared savings) that reward year-over-year improvement and *attainment incentives* that reward continued excellence in resource stewardship relative to population benchmarks. The incentives are then combined; POs that both improve and meet attainment benchmarks are eligible to earn both incentives (see Exhibit 1).

In considering how to structure a shared-savings program, one of the most important decisions was whether to base the incentive only on upside risk or include downside risk as well, which would expose physician organizations to financial losses. Of the approximately 200 participating POs in MY 2014, 49 percent reduced utilization sufficiently to generate net savings and earned a share of those savings. Another 16 percent did not pass one of the two eligibility gates—quality and total cost of care trend. The remaining 35 percent of POs experienced net losses but were not financially responsible for sharing in those losses because the program does not include downside risk. Exhibit 2 shows the distribution of net savings and losses among POs for MY 2014.

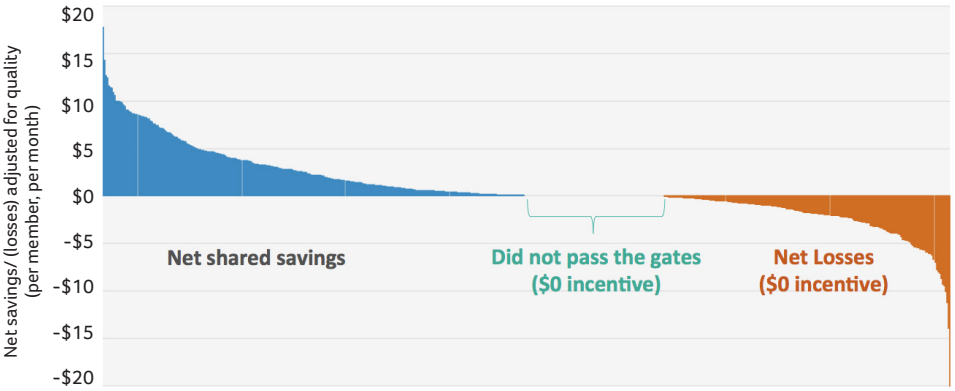
Upside-only risk keeps the dynamic entirely positive and is central to provider engagement, which is especially

Exhibit 1: IHA Value Based P4P Incentive Pathways



Source: Integrated Healthcare Association Value Based P4P Program.

Exhibit 2: Net Physician Organization Shared Savings and Losses, MY 2014



Source: Integrated Healthcare Association Value Based P4P Program.

important in the early development and buy-in phases of a program. Incorporating downside risk into the design can strengthen the incentive and create a larger funding pool since losses among some providers fund gains among others without requiring additional dollars from program sponsors, typically health plans or other payers.

For the IHA Value Based P4P program, where participation is voluntary and there was an existing, established budget for incentives, the decision was to go with an upside-only design, partially mitigated by requiring *net* improvement where PO losses can be offset by gains. If the net is not positive, the PO receives no shared savings. For example, additional costs associated with an increase in readmissions can be offset by savings generated from reducing emergency department (ED) visits. This approach represented a substantial shift from previous health plan incentives for utilization, where categories were generally considered separately—so a bonus payment could be earned for lower ED visits even if readmissions increased. The shift to a net improvement approach aligns with goals of broader accountability and can be viewed as a step toward downside risk.

Design Decision No. 2: Total Cost of Care vs. Utilization as Basis for Incentive Payments

A major focus in transitioning from quality-specific to a broader value focus for the program was identifying measures of cost to incorporate into the value equation. The program already had a well-established and accepted definition of quality developed through participant consensus, and that definition served as the "Q" in the value equation. The challenge, accordingly, was measuring cost; the key candidates for that side of the value equation are resource use (utilization) measures and total cost of care.

While utilization metrics are frequently tracked and monitored by both health plans and providers, less effort has been invested in refining and standardizing performance measures based on utilization as compared to quality. Efforts to create standardized performance measures typically narrow definitions to certain populations. In quality performance assessment, measures are carefully specified—the relevant population segment is identified to ensure consistency and comparability and to clarify the performance signal. In a shared-savings environment, by contrast, every additional slice and stratification diminishes savings opportunity and creates a potential disconnect between the measured performance and the observed outcomes, which matter a great deal from the payer and purchaser perspective.

The two main candidates for incorporating cost into the incentive design—utilization measures and total cost of care—pose tradeoffs that the decision-making committees discussed extensively (see Exhibit 3 for a side-by-side comparison of the approaches), including:

Aim. How strong is the link between the performance measure and the desired outcome? At a system level, the goal is higher value—better quality at lower or comparable costs. Total cost of care is directly tied to value, while utilization is an intermediate step. In fact, because cost is the result of both utilization and price, reductions in utilization may be insufficient to drive cost reductions.

Control. How much control do physician organizations have over what is being measured? Pricing—facility and drug—is a key factor in total cost, but POs often have little price information to factor into referral decisions because of the confidentiality of negotiated rates between plans and hospitals. Without both price and quality information, it is impossible to refer patients to the highest-performing specialists and facilities from both quality and cost standpoints. In contrast, POs have much more control over such decisions as whether a patient is admitted to a facility at all or prescribing brand vs. generic drugs.

Without both price and quality information, it is impossible to refer patients to the highest-performing specialists and facilities from both quality and cost standpoints.

Engagement. Are providers motivated to perform well on the measure? Most providers believe that strong primary care coordination has the potential to reduce admissions, ED visits, and readmissions. Improving patient care and quality is highly motivating from a clinician perspective, and reducing unnecessary utilization has a beneficial side effect in lowering costs. By contrast, the goal of reducing total cost of care is less compelling to clinicians.

Feasibility. How difficult is it to get accurate results? From a technical standpoint, tracking total cost of care is much simpler and less costly than relying on resource use measures for a variety of reasons. Hospitalizations and readmissions are relatively rare events with very low incidence in certain populations. In contrast, total cost applies broadly to the population, so results are more reliable, especially in the capitated/delegated

HMO model. Cost data tends to be more accurate and complete because it drives payment, while utilization data quality varies and may degrade with multiple handoffs of encounter and claims data. Risk-adjustment tools are readily calibrated to predict total cost of care, while fewer tools are available to risk adjust utilization measures and/or risk adjustment methods for utilization need to be derived.

While early Value Based P4P design discussions favored total cost of care as the basis for shared savings, ultimately, the decision to use utilization measures was driven by the need to engage POs in a meaningful way. California POs also were more comfortable with utilization—they had been receiving utilization reports from health plans for years, representing similar metrics such as bed days and ED visits. And, the utilization measures for P4P were tested and implemented before the TCC measure, so there was more familiarity with the results. Further, there was a strong concern about the lack of underlying provider price information, which was a significant hurdle to gaining PO buy-in for the TCC measure. And, without sufficient breakdown of various cost components, the monolithic nature of a total cost measure makes identifying the source of a cost problem and possible solutions an additional challenge.

The rationale for selecting utilization as the basis for incentive payments at the time was clear, but the pendulum is now swinging back toward total cost of care. TCC is in use for awards and public reporting, prompting renewed interest in reconsidering the role of TCC in the Value Based P4P program. As value-based initiatives expand across the health care landscape, the acceptability of considering cost performance has increased dramatically. The National Quality Forum (NQF) endorsement of HealthPartners' total cost of care measure in 2012 is one example of the spread of affordability into the historically quality-based performance measurement arena.

Exhibit 3: Utilization vs. Total Cost of Care as Basis for Incentives

Consideration	WHICH APPROACH SCORES BEST? Utilization (Resource Use)	Total Cost of Care
Aim: Link to desired outcome of lower costs	Lower—utilization is reasonable proxy for cost, but each measure is only a slice of the total and price is excluded	Higher—direct link to goal of lower costs
Control: Extent to which POs control the outcome	Higher—providers order services on behalf of patients	Lower—prices are key element of total cost of care but often outside provider control and unknown
Engagement: Actionability and sensitivity of data	Higher—established area of health plan reporting; logical link between high-quality care (which motivates clinicians) and lower admissions and readmissions (with a beneficial side effect of lower costs)	Lower—negotiated rates are confidential with legal barriers to sharing; clinicians are inspired by providing great care not by lowering costs
Feasibility: Availability of valid measures and risk adjustment	Lower—available measures have less consistent specifications, and risk adjustment tools generally calibrated to cost rather than utilization	Higher—cost measures have greater consistency in specifications, and an National Quality Forum-endorsed measure is available; standardized risk-adjustment tools are available (though often proprietary and costly)

Source: Integrated Healthcare Association Value Based P4P program.

Among IHA stakeholders, rollout of physician organization awards and the release of TCC star ratings for public reporting generated new attention to cost measurement and results. TCC presents a singular and compelling reflection of affordability that can be easily compared against composites of quality and patient experience—potentially supporting not just provider improvement but also health plan development of provider networks, purchaser offerings, and consumer choice. All of these factors led to the recent decision to transition, over the next several years, from utilization to TCC as the basis for incentive payments.

Design Decision No. 3: External vs. Participant Population Benchmarks

Development of any incentive design requires benchmarks to assess performance. In the many and lengthy discussions among Value Based P4P stakeholders about benchmarks, external benchmarks initially gained favor. Benchmarks using comparison data that is separate and distinct from participants' experience have natural appeal: The independence of the benchmarks from the measured population prevents rewarding performance improvement that is part of—but no better than—secular trends. That is, if

a widespread shift in practice patterns across the country results in a systematic increase in quality—for example, reduction in antibiotic use for acute bronchitis—that may not represent strong performance as much as standard practice.

In the real world, however, external benchmarks are frequently impracti-cal—they may not be available, timely, comparable, or applicable to the need. While benchmarks are available for Healthcare Effectiveness Data and Information Set (HEDIS) measures, they may not be available for other measures, such as the Consumer Assessment of Healthcare Providers and Systems Clinicians & Groups Survey. Another obstacle is that specifications may be proprietary, making it impossible to assess comparability; examples include benchmarks produced by Milliman, MedInsight, and Verisk. Reporting lags with published benchmarks frequently make them unusable, since current data cannot be compared to benchmarks from several years prior. Specifications can change frequently, invalidating benchmarks based on prior specifications. Additionally, industry trends can cause breaks in benchmark continuity as the health care environment changes. Patent expirations can affect benchmarks for generic prescription measures even if prescribing practices don't change, and the transition to ICD-10 coding resulted in changes to multiple benchmarks based on ICD-9 codes.

Identifying external benchmarks for measures of cost and utilization is particularly challenging because there are relatively few available, and those that exist tend to lack consistency of specifications and risk-adjustment methodology. A measure of bed days is commonly used, but handling of days for maternity, mental health, or the length of claims runout may have considerable impact on observed rates. At the time the total cost of care measure now used in Value Based P4P was developed, there were not any NQF-endorsed measures for health care costs.

Exhibit 4: California and National HEDIS Utilization for HMO/POS – Health Plan Average, 2014

Measure Name	National	California
Ambulatory Care: Emergency Department Visits (PTMY)	194.9	144.4
Inpatient Utilization: Total Inpatient Days (PTMY)	180.1	135.5

Note: PTMY = per thousand member years.
Source: National Committee for Quality Assurance Quality Compass® 2015 (reflects performance in 2014).

Even when benchmarks are available, they may not be relevant. California's substantially lower utilization often invalidates comparison to national benchmarks—for example, ED visits per thousand member years is 144.4 for California HMO and 194.9 for national HMO, according to HEDIS MY 2014 data (see Exhibit 4).

While using benchmarks based on the participants' patient population may seem like a compromise of necessity, they are often more meaningful—and therefore credible—from a provider perspective. Benchmarks based on 'peer' performance represent attainable—though perhaps challenging—targets.

Since then, HealthPartners' Total Cost of Care (TCOC) Index has received endorsement, which IHA strongly supported. The endorsed measure is similar to the IHA VBP4P TCC measure; both measures are risk adjusted and reflect health care costs across providers and settings. However, a few important differences restrict the comparability of the results and, therefore, of benchmarks. For example, the IHA TCC measure includes all ages, while the HealthPartners TCOC measure excludes infants under 12 months of age and adults older than 65. Risk adjustment presents a related challenge. While the performance of the various risk-adjustment models for cost yields comparable rankings and explanations of observed variance, the results from different risk-adjustment models are not directly comparable.[2]

Benchmarks based on the participant patient population present an alternative to using external benchmarks. While using benchmarks based on the participants' patient population may seem like a compromise of necessity, they are often more meaningful—and therefore credible—from a provider perspective. Benchmarks based on "peer" performance represent attainable—though perhaps challenging—targets. Given the many challenges associated with using external benchmarks, it is not surprising that internal benchmarks dominate in the Value Based P4P program. After careful consideration, only one out of the six benchmarks in the final design—TCC trend—incorporated

Exhibit 5: IHA Value Based P4P Benchmarks

Design Element	Benchmark	External	Notes/Rationale
Quality Gate	Quality composite score of 10 points or higher, which means, on average across measures, the physician organization (PO) either 1) met the 75th percentile threshold, or 2) closed the gap 10% between its prior year performance and the benchmark	No	• Emphasizes the importance of quality in the program, and the threshold can be met through a combination of attainment and improvement • Exhibit 6 shows the results when the gates are applied to the PO population
TCC Trend Gate	**Standard Threshold** MY 2013-2015: CPI+3 percentage points MY 2016-2017: CPI+2 percentage points MY 2018-2019: CPI+1 percentage point **High-Cost POs** MY 2013-2015: CPI+1 percentage point MY 2016-2017: CPI+0 percentage points MY 2018-2019: CPI-1 percentage point	Yes	• The recommendation is to use a three-year average of the U.S. Consumer Price Index, which would be based on the measurement year and the two years immediately preceding the measurement year • The total cost of care (TCC) trend gate for high-cost POs is set at two percentage points below the standard threshold • Exhibit 6 shows the results when the gates are applied to the PO population.
High-Cost PO Definition	POs with geography- and risk- adjusted TCC above the plan-specific 90th percentile for both baseline and measurement year	No	• The VBP4P committees believe it is important to hold POs that have consistently high costs to a stricter TCC trend gate to strengthen incentives for cost reduction
Utilization Shared Savings	Own prior year performance	No	• After considering adjustments, the VBP4P committees opted for own-organization performance because it is easier to understand—an essential ingredient for PO engagement
Utilization Attainment	75th percentile for participating population for both current and prior year (see Exhibit 7)	No	• The VBP4P committees agreed that **maintaining** relatively low utilization merits reward, even if it doesn't generate new year-over-year savings
Quality Adjustment	The recommended value is set to provide about a two-fold difference between the lowest qualifying performers and the highest performers; multipliers range from 0.65 to 1.35, which correspond with a PO earning a share of savings between 32.5% and 67.5%	No	• Adjusting the savings for quality reinforces the importance of quality in the VBP4P design

Source: Integrated Healthcare Association Value Based P4P program.

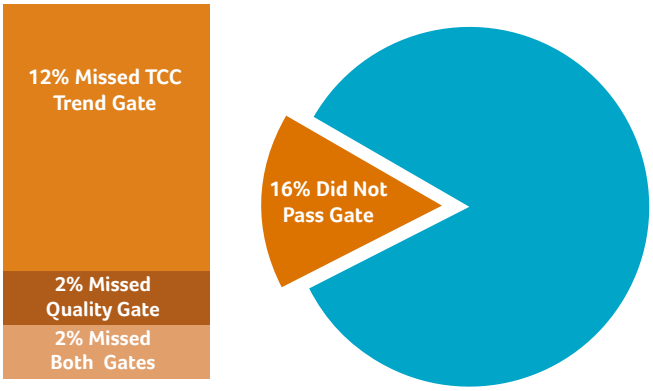
an external benchmark (see Exhibit 5).

As shown in Exhibit 6, when the gates for quality and total cost of care are applied to the full population of participating POs for MY 2014, 16 percent do not reach the thresholds and are ineligible for shared savings. Exceeding the total cost of care trend gate is the main reason, with 12 percent of POs ineligible; an additional 2 percent did not pass the quality gate, and 2 percent did not reach either gate.

The utilization benchmarks for the attainment incentives are shown in Exhibit 7. The benchmarks are set at the 75th percentile of results from the participating PO population for both the baseline and measurement years. For example, POs that met the attainment threshold on readmissions for both years—6.1 percent for MY 2014 and 6.4 percent for MY 2013—would earn a supplement to any shared savings generated through reductions in utilization.

Regardless of whether the benchmark is drawn from external sources or the participant population, an ongoing challenge, as previously mentioned, is balancing simplicity with methodological rigor. The Value Based P4P program is undeniably complex. Simpler approaches to incentive design are more understandable, making it easier for participants to see how their

Exhibit 6: Applying Quality and Total Cost of Care (TCC) Trend Gates to Physician Organization Scores for MY 2014



Source: Integrated Healthcare Association Value Based P4P program.

Exhibit 7: Benchmarks for Value Based P4P Utilization Attainment Incentive, MY 2014

	Resource Use Measure	Measurement Year (MY 2014) 75th percentile	Baseline Year (MY 2013) 75th percentile
↓ Lower is better	All-Cause Readmissions	6.1%	6.4%
↓ Lower is better	Inpatient Bed Days (PTMY)	84.5	94.0
↓ Lower is better	Emergency Department Visits (PTMY)	139.6	137.0
↑ Higher is better	Generic Prescribing (overall)	86.7%	84.8%

Note: PTMY = per thousand member years.

Source: Integrated Healthcare Association Value Based P4P program.

actions lead to program outcomes. Yet, to enable broad participation and maximize credibility of results, modifications are often worth the increased complexity that they bring. For example, rather than excluding small physician organizations because their results are less reliable, the program incorporates an adjustment to the utilization results that adds complexity but broadens participation.

Design Decision No. 4: Investing in Performance Improvement

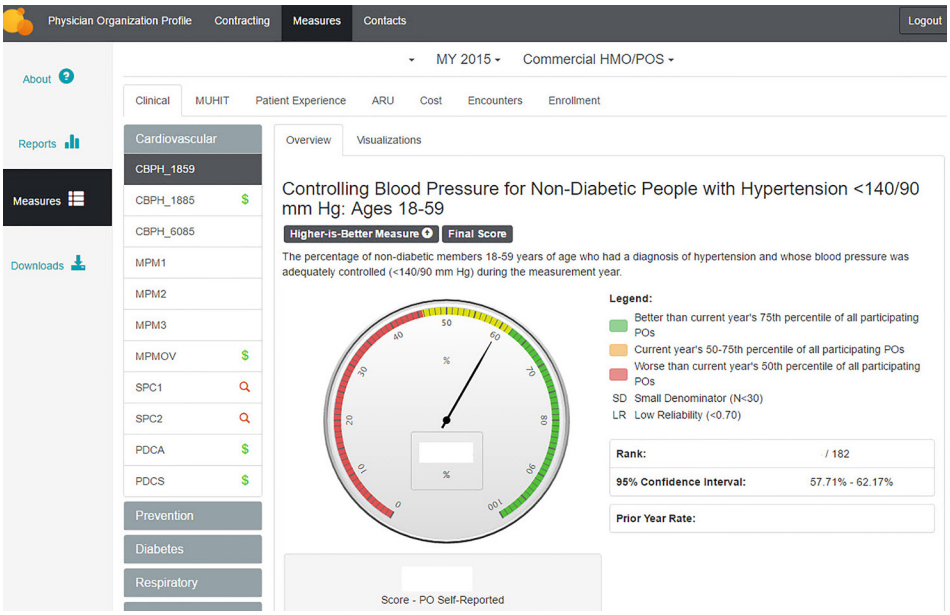
Engaging in performance measurement and reporting generally requires substantial resource investments—data systems and registries, staff to support quality improvement initiatives, preparation of data for reporting, and collection of survey-based results. As a result, it is critical that these investments are worthwhile to P4P participants. An important component in the transition from a quality-only to value-based incentive design has been the development of tools that participating physician organizations and health plans can use to help them understand and use the results to improve performance—the ultimate goal of any P4P program.

When the program was focused specifically on quality, participating medical groups and IPAs could rely largely on information and systems available within their organizations. The transition to incorporating utilization and total cost of care measures into the design introduced the need for information outside of the physician organization's immediate purview—information on prescription drugs and inpatient bed days, for example, that comes from health plans. Broadening data sources that contribute to performance results is consistent with the broader movement toward greater accountability but brings substantial challenges; utilization results are inherently complex and multifaceted. While performance improvement is never easy, identifying causes and interventions for condition-specific measures—such as a low breast cancer screening rate—is much more straightforward than reducing bed days, which have many and varied underlying causes.

To support participants' ability to understand the results and make them more actionable, the program has made significant investments in developing modeling and reporting tools.

Reporting Portal. The VBP4P reporting portal (see Exhibit 8) is the platform used to distribute data and reports to PO and health plan participants. While the view differs for POs and health plans, the reporting portal visualizes and communicates performance on quality, utilization, and cost metrics collected in Value Based P4P. These visualizations are metric specific to be easily and widely understandable to participants. POs see performance for each plan they contract with, as well as in aggregate; health plans see PO-specific breakdowns for POs they contract with, as well as in aggregate. Additionally, underlying measure-specific data, aggregated

Exhibit 8: Screenshot of IHA Value Based P4P Reporting Portal



Source: Integrated Healthcare Association Value Based P4P program.

data, and thresholds are made available as .csv downloads to participants that would like to manipulate and visualize their own data.

Interactive Modeling Tools. The program provides spreadsheet-based modeling tools to all participants to support greater understanding of the VBP4P design (see Exhibit 9). The tools allow both POs and health plans to assess the impact of "moving the dial" on each design component.

- The PO worksheets depict the core Value Based P4P design using a physician organization's own performance on quality, resource use, and total cost of care for the current measurement year. Worksheet cells can be dynamically changed to understand the impacts of various performance levels on health plan-specific incentives. The calculations are not intended to reflect the actual payments received from participating health plans; rather, they are provided for illustration of the general design.

Exhibit 9: Screenshot of IHA Value Based P4P Modeling Tool for Physician Organizations

Value Based P4P Incentive Worksheet - Measurement Year 2015
PHYSICIAN ORGANIZATION (00000000)
HEALTH PLAN

Value Based P4P Incentive Design for Shared Risk POs

Performance Gates				
Gate	**Measure**	**Threshold**	**Your Results**	**Pass (Y/N) ?**
Quality	Quality Composite Score[2]	10	39	Yes
Total Cost of Care (TCC) Trend[4]	TCC Trend with 85% CI	4.1%	-2.5%	Yes
Total Cost of Care (TCC) Amount	MY2014 TCC ($PMPY)	90th Percentile	< 85th	Yes
	MY2015 TCC ($PMPY)		90th to 94th	
			Quality Multiplier	0.96

Shared Savings based on Appropriate Resource Use Measures

	Measures	MY 2015 Result [1]	MY 2014 Result [1]	Change in Rate	Normalized Expected Rate[6] (MY 2015)	Denominator	Denominator Units	Resource Use Saved/ (Increased)	Cost per Resource Use	Estimated Savings/ (Additional Cost)
	a	b	c	d = Δ(b,c)	e	f	g	h = d*e*f	i	j=h*i
↓	All-Cause Readmissions (PCR)	1.65	0.44	(1.21)	12.5	35	Admissions	(5.3)	$14,000.00	($73,811.96)
↓	Inpatient Bed Days (IPBD)	78.90	87.33	8.43	NA	3,013	1,000 Member Years	25.4	$4,000.00	$101,555.41
↓	Emergency Department Visits (EDV)	166.19	165.67	(0.52)	NA	3,013	1,000 Member Years	(1.6)	$750.00	($1,185.95)
↑	Outpatient Procedures in Pref. Facility (OSU)	124.00	102.10	21.90	NA	40	O/P procedures	8.8	$1,500.00	$13,140.00
↑	Generic Prescribing (GRX) - Overall	89.12	87.52	1.60	NA	18,245	# Prescriptions	291.9	$70.00	$20,434.40

LEGEND
Health Plan Values *- Your PO can update these cells to reflect your own assumptions and understanding of health plan VBP4P incentive programs.*

Net Estimated Savings/Additional Cost	$60,132
Net Savings? (Y/N)	Yes
PO Share of Net Savings or (50%)	$30,066

Source: Integrated Healthcare Association Value Based P4P program.

- The health plan "Plug & Play" tool summarizes performance for all POs that have active contracts in place with that health plan for the current measurement year. It is intended to assist health plans in determining which POs are eligible for an incentive and to model what incentives are earned under the VBP4P incentive design.

Member- and Claims-Level Detail Files. Generated from health plan claims submissions for resource use and total cost of care measurement, these files are created by Truven Health Analytics and delivered to health plans to distribute to POs. IHA does not handle these files, which contain protected health information. These files allow users to drill-down to member-level data.

Design Decision No. 5: Fostering Broad and Public Use of Results

While the opportunity for shared savings is an important incentive, it certainly is not the only incentive in Value Based P4P. From the program's inception, the performance measure results were used for public reporting and PO recognition. To signal the change in philosophy that quality alone is no longer sufficient, it was critical to incorporate cost into these nonfinancial incentives.

IHA partners with the California Office of the Patient Advocate to create the Medical Group Report Card, which is the largest statewide multipayer public report card to provide side-by-side assessments of physician organization performance on all three key aspects of value: clinical quality, patient experience, and costs. A four-star rating system shows overall performance in each of these three areas for all physician organizations serving a selected county (see Exhibit 10).

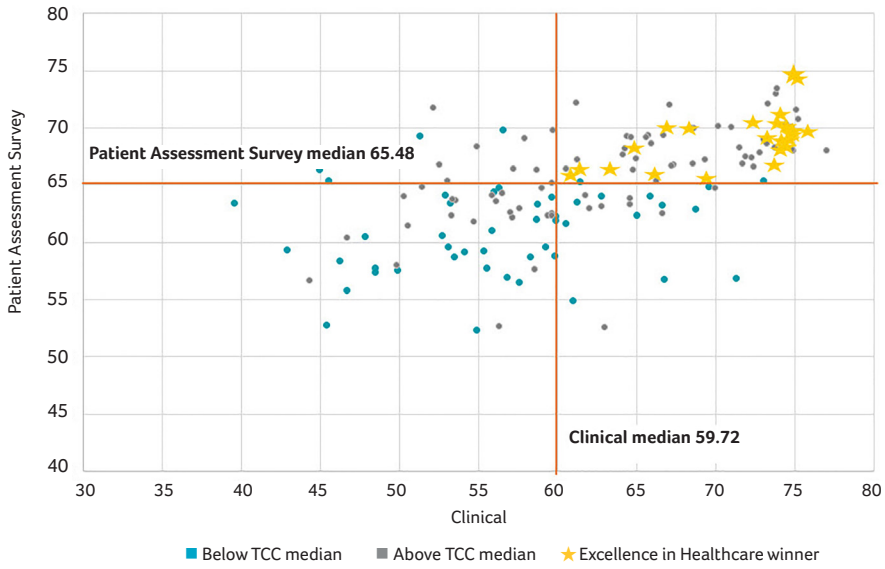
Clinical quality and patient experience have been publicly reported for over a decade. Agreeing on a common way of measuring total cost of care and gaining buy-in to report the results publicly took years of groundwork by IHA and participating physician organizations and health plans; it came to fruition when total cost of care was added to the report card in early 2016. Physician organizations can now see how they compare with one another and, ideally, compete on quality and cost; consumers now have transparent, audited quality and cost information to consider when selecting a physician organization; purchasers and health plans have more complete performance information about their contracting partners; and policymakers can identify care delivery and cost patterns across the state.

In 2014, the first Excellence in Healthcare award was presented to top-performing physician organizations. To earn the award, physician organizations must perform in the top 50 percent for clinical quality *and* patient experience *and* total cost of care. Out of more than 200 participating physician organizations, only 23 met the standards in MY 2014—the stars in Exhibit 11. The

Exhibit 10: Public Report at the Physician Organization Level, Office of the Patient Advocate

Source: Integrated Healthcare Association Value Based P4P program.

Exhibit 11: Physician Organization Excellence in Healthcare Award Winners, MY 2014



Source: Integrated Healthcare Association Value Based P4P program.

program retired the quality-only award in 2015—a milestone signifying that physicians can no longer be classified as "top performing" without doing well on cost as well as quality.

Implications

At the national level, public and private payers increasingly are moving toward value-based payment methods that begin to align incentives for providers to deliver high-quality, affordable, patient-centered care.[3] Gaining critical mass across public and private payers in holding providers accountable for both the quality and cost of care through aligned performance measures, incentives, and public reporting is essential to scaling the triple aim of better care, better health, and smarter spending across the delivery system. Health care markets differ, and stakeholders must work through how best to design strong and consistent signals that advance high-value care based on local priorities and market and participant characteristics. Nonetheless, the IHA Value Based P4P program experience to date can be instructive on some of the most important—and challenging—design decisions.

In California, the transition from allocating "bonus" funds—P4P dollars budgeted by health plans—based solely on quality to a value-based focus that includes both quality and costs and is funded by shared savings represented a sea change for participating physician organizations, which arguably are among the most integrated in the country. As the P4P program transitioned to a value-based focus, where strong performance on quality doesn't necessarily translate into high performance on value, there have been some difficult—but ultimately constructive—discussions among participants.

The core components of the IHA Value Based P4P design are relevant beyond the commercial HMO product line, and the program doubtless will continue to evolve as California health care markets continue to change. For example, medical groups and IPAs in the state are entering new product lines, such as Covered California (the state health insurance exchange) and Medi-Cal (California Medicaid), and experimenting with new risk arrangements such as accountable care organization contracts. At the same time, the IHA Value Based P4P program continues to learn from participants and to support their performance improvement efforts as we all share navigation on the expedition to high-value care.

Notes

1. Simon, L.P., and A. Monroe, "California Provider Group Report Cards: What Do They Tell Us?" *American Journal of Medical Quality* (March/April 2001).
2. Winkelman, R., and S. Mehmud, "A Comparative Analysis of Claims-Based Tools for Health Risk Assessment," Society of Actuaries (April 2007). Online at https://www.soa.org.
3. See, for example, the Health Care Transformation Task Force at http://www.hcttf.org/aboutus/ and Health Care Payment Learning & Action Network at https://hcp-lan.org/.

Acknowledgments

This project was generously supported by a grant from the California Health Care Foundation (CHCF). Ginamarie Gianandrea and Thien Nguyen contributed to the development of the issue brief, and Alwyn Cassil of Policy Translation, LLC, provided invaluable editing services. We are also grateful to reviewers from health plans, physician organizations, and CHCF who provided helpful feedback. The Value Based P4P program operates under the auspices of the Governance Committee, Technical Measurement Committee, and Technical Payment Committee; committee members provided strategic guidance and ultimately made the key design decisions described in this issue brief during the transition from quality to value.