How-To Guide: Evaluating Payment Reform Programs

December 2015

How to Use this Guide

This guide serves as an introduction to **CPR's** **Payment Reform Evaluation Framework** and is intended for use primarily by self-funded employers and other purchasers of health care interested in evaluating the outcomes of a particular payment reform program. We hope it will also be useful to health plans, health care providers and others in the health care system.

The Payment Reform Evaluation Framework has four core domains, each of which contains a targeted list of questions aimed at the sponsor of the payment reform program, such as a health plan, to assess whether the program is having a positive impact and would be appropriate or beneficial to the purchaser's population. The four domains of the Framework include questions seeking a general description of the program, and those specifically related to the program's feasibility, cost, and quality.

The Framework can be used to evaluate an existing program or to design the evaluation of a new program. It is meant to be used to evaluate one program at a time, not an entire payment reform strategy. It can also help purchasers assess the business case for involving the population for whom they purchase health care in a program.

Ultimately, we hope the Framework helps establish standard parameters for evaluation in the marketplace, enabling the comparison of results from a variety of different programs, the identification of successful approaches and the further spread of those approaches.

I. Introduction

As health care costs for employers and other health care purchasers and consumers continue to rise, without solid evidence that the quality of health care is also improving, purchasers, health plans and health care providers have begun to experiment with delivery system and payment reform programs. The goal of such programs is to decrease overall health care spending, while improving the quality of care. However, while many payment reform programs have been implemented over the past few years, limited data exist to demonstrate their effectiveness; it is still largely unclear whether they will produce their intended outcomes of better quality and lower costs.

This uncertainty regarding outcomes is due, in part, to the lack of a standard evaluation framework to help gauge the impact of such programs. As a result, every evaluation is different. At best, this means evaluations are well suited to provide results for specific programs, but these results cannot be compared across programs. At worst, this gap allows program sponsors to tailor their evaluations to highlight successes and deemphasize failures.

Therefore, the objectives of this guide are to 1) introduce purchasers to a standard approach for evaluating payment reform programs, 2) help purchasers identify which metrics might best help them to evaluate a payment reform program's impact on health care quality and cost, 3) send a unified message to providers and program sponsors as to what outcome measures are most important to employers and other purchasers of care, and 4) influence others in a position to evaluate payment reform programs, such as health plans, health care providers, and benefit consulting firms, to encourage rigor and consistency when performing analyses of the outcomes.

The Payment Reform Evaluation Framework is meant to be used on one payment reform program at a time – it is not meant to evaluate an entire, multifaceted payment reform strategy. For example, a health plan might administer a prospective bundled payment program with 10 of its contracted orthopedic practices for hip and knee replacements. This framework could be used to evaluate the specific prospective bundled payment program.

In addition to standardizing the evaluation process, the Payment Reform Evaluation Framework can help purchasers considering pursuing a new payment reform program to make a clear and measurable business case for doing so. The Framework provides answers to four critical questions commonly asked when "making the case" for program adoption:

1. Did the program reduce health care spending in relationship to the trend, or at least keep the trend flat?
2. Did quality of care improve, or at least stay the same?
3. Was consumer feedback positive, or at least neutral?
4. Is the program feasible to implement, replicate, scale, and maintain over time?

While payment reform programs are a popular and potentially effective way to reduce costs and increase the quality of care, a standard evaluation process is essential to ensuring that these programs achieve their intended outcomes and create value.

What are Payment Reform Programs?
For the purposes of this guide and the Payment Reform Evaluation Framework to which it refers, a health care payment reform program is a discrete arrangement between a health care purchaser or payer and a health care provider or providers that utilizes an alternative, value-oriented method of provider payment to drive both improvements in the quality of care and reductions in the cost of care for a defined population.

Today, the vast majority of payment reform programs utilize one or more of the following payment methods:

- Fee for service (FFS) plus shared savings and shared risk, such as those arrangements that support accountable care organizations (ACOs), patient centered medical homes (PCMH)
- FFS plus shared savings only
- Capitation (full, partial, or condition-specific) in which provider payment is based, in part, on quality performance
- Bundled payment in which provider payment is based, in part, on quality performance
- Quality incentives (e.g. FFS plus pay for performance)
- Payment for non-visit functions (e.g. care coordination fees, investments in health information technology)
- Non-payment policies for specific services or events (e.g. hospital acquired conditions or early elective deliveries)

More detailed definitions of these payment methods are available here.

The Need to Evaluate Payment Reform Programs
An enormous movement is underway to change how providers are paid in both the public and private sectors. This investment in time and effort will be well worth it if these reforms result in significant improvements in the quality and affordability of care, while reducing the overall cost of care. Currently, however, the research that is available is mixed; while one study shows a particular payment method to be effective, the next one questions its efficacy. Therefore, it is essential to experiment, evaluate results, and continue to innovate over time.

Complicating matters is the fact that program results may depend heavily on context. How well a given payment reform program succeeds depends on the skills and motivation of the specific parties involved, the design of its incentives, the implementation approach, the population it serves and the dynamics of the market in which it is implemented.

For example, a particular payment reform program may be far more successful if the majority of health care purchasers participate, thereby making the incentives in the program more meaningful to participating health care providers. However, the results of one program may not be generalizable. Thus, there is a pressing need to amass more evidence on the impact of various approaches to be able to draw general conclusions, and it will be far easier to aggregate these results if we use a standard approach. A consistent method of evaluation will also allow for greater comparison of outcomes across programs, and send a clear signal to program sponsors about what types of impact are most meaningful to purchasers and other stakeholders.

In addition to the methodological benefits of a standardized evaluation process, it may also produce richer data. Program sponsors and others being asked to provide information may be more willing to do so if they aren't receiving multiple requests from multiple parties, especially when each request seeks information that is slightly different from the next. The reporting burden on sponsors, plans, and providers has been a significant issue in the quality measurement movement, and it has resulted in resistance to using meaningful measures.

Such evaluation will also address key, practical questions such as: Should an existing program continue? Are changes to the program needed to improve results? Is it feasible to replicate the program or expand its scale? Employers and other health care purchasers will want to know the answers to these questions, and hold program partners accountable for results they have promised or are contractually obligated to produce.

How Should We Evaluate Payment Reform Programs?

There are many approaches to evaluating payment reform programs and there is no "one size fits all" method; however, standardization could be extremely beneficial to the industry, for the reasons described above. Recognizing this, CPR sought input from an expert advisory committee to develop a framework to evaluate payment reform programs. Our advisory committee included:

- Michael Bailit, Bailit Health Purchasing
- Michael Chernew , Harvard University
- David Cowling, CalPERS
- David Cutler, Harvard University
- Guy D'Andrea, Discern Health
- Brooks Daverman, Div. of Healthcare Finance and Administration, State of TN
- François de Brantes, Health Care Incentives Improvement Institute
- Vicky Ducworth, The Boeing Company
- Anna Fallieras, GE
- Robert S. Galvin, Equity Healthcare
- Paul Ginsburg, University of Southern California
- Mark McClellan, The Brookings Institution
- Elizabeth Mitchell, Network for Regional Health Improvement
- Dana Safran, BCBS of Massachusetts

In addition to the substantive input from the advisory committee, we obtained feedback on the feasibility of the Framework from three national health insurers. CPR also sought input from its diverse member base, which includes large private employers, and state Medicaid, employee and retiree agencies. We would like to thank all participants in this process for their contributions, though CPR takes full responsibility for any shortcomings.

The Advisory Committee emphasized the need to focus the evaluation on the cost and quality impacts of a payment reform program. CPR also felt strongly about assessing the feasibility of a program, as the value of any cost or quality gains can be amplified if the program can be sustained, scaled, and replicated. The health plans generally expressed confidence that they could respond to the metrics, but voiced concerns about the reporting burden they face, given that many different parties are asking them to share and report results. Lastly, CPR employer-purchaser members noted the importance of staging the evaluation process, starting with gathering background information to put results in context, and then collecting outcomes information at regular intervals over time.

To be clear, we did not focus on research methodology. We did not determine, for example, whether it is always best to have a comparison group or to use particular benchmarks, or whether an evaluation is only credible if it contains a randomized experiment versus a difference-in-difference analysis. Instead, we concentrated on the domains we believe are most critical to address in any approach to the research, including program design, feasibility, cost and quality, and the specific elements within those domains employers and other health care purchasers most want to know about.

Overview of CPR's Payment Reform Evaluation Framework

CPR's Framework is divided temporally into two sections:

1) Initial Assessment, focused on the four domains of program design, including a general description of the program, and questions related to the program's feasibility, cost, and quality.
2) Ongoing Monitoring, focused on program outcomes regarding cost and quality.

1. Initial Assessment

Program Design: General Description

This section is meant to gather largely descriptive, background information from the program sponsor about the nature of the program, including its goals, the lines of business in which it is offered, its availability by region and market, which purchasers are eligible to participate, and for which insurance products it is available. The section also asks about the provider payment methods at play, how quality targets are set, and what complementary benefit designs are in place or could enhance the success of the program. The questions in the General Description section are meant to provide context for any results, as well as give the prospective participating purchaser a sense of what to expect.

Program Design: Feasibility
The Feasibility domain tries to gauge whether the program is viable from an implementation standpoint, asking questions related to additional administrative costs or investments that may be required on the part of various stakeholders. This section also inquires as to whether the program can be replicated and scaled, recognizing that programs that succeed in a singular situation, but are difficult to implement, will have limited application. The Feasibility section also asks what means are used to ensure provider and consumer participation in the long term.

Program Design: Cost
The Cost domain seeks details on the mechanisms through which the program intends to reduce costs, its impact on total health care spending in comparison to benchmarks, and how any savings may be passed on to employers and other health care purchasers. The Cost section also features questions for specific payment models including bundled or episode-based payment, shared savings and shared risk, capitation, and pay-for-performance.

Program Design: Quality
The Quality domain focuses on various clinical quality and patient satisfaction and experience measures. This section also inquires as to whether the program ties provider payments to important quality indicators, incentivizes providers to improve their performance based on these quality indicators, and/or uses quality measures to evaluate results.

Unique to this domain of the Framework, this section references another CPR resource, the CPR Employer-Purchaser Priority Measure Set. In doing so, the Quality section seeks to assess whether the payment reform program addresses any of the twelve priority clinical areas highlighted in the Priority Measure Set. CPR chose these clinical areas based on an HCI^3 analysis of commercial claims data to identify conditions representing the biggest area of spend for which there was also large variation in the quality of care and payment amounts. The 12 priority clinical areas HCI^3 identified are: pregnancy, hypertension, low back pain, diabetes, depression, osteoarthritis, breast cancer, arrhythmia, asthma, coronary artery disease, gastrointestinal endoscopy, and upper respiratory infection. Discern Health helped CPR identify the best measures to pair with these priority clinical areas, as well as "cross cutting quality measures" that gauge person-centeredness, preventive care, and patient safety, emphasizing National Quality Forum-endorsed measures and outcomes measures wherever possible.

2. Ongoing Monitoring

Program Outcomes: Cost and Quality
These sections are designed to help purchasers collect outcomes data for the program on an ongoing basis. The Cost section focuses on whether the program generates savings, incurs costs, or has an impact on total health care spending. This section also requests program sponsors to report any changes in performance on measures of efficiency. The Quality section asks how the program has impacted performance on quality measures that could identify any unintended negative consequences resulting from incentives created by the program, as well as

the results on the clinical quality and patient satisfaction and experience measures the program uses.

When evaluating payment reform programs, even while leveraging a standard approach, it is important to acknowledge that there will always be challenges to generalizing results. Market dynamics are likely to impact the success of programs; for example, in markets with significant competition among health care providers, providers may be more likely to accept and succeed under payment methods that bring them financial risk. Benefit design has an impact as well – do consumers have an incentive to seek providers who offer higher-value care, and will that patient volume be enough to convince providers to accept and succeed under new forms of payment? There are so many factors at play in every payment reform program that it may not be possible to draw direct causal lines between the reform approach and results. However, if applied consistently, the Payment Reform Evaluation Framework could help purchasers and other stakeholders compare the results of various programs offered by a common sponsor, as well as compare the outcomes of similar programs across different program sponsors, and amass results over time.

Who should use the Payment Reform Evaluation Framework?

Many different parties may be in a position to use this Framework for program evaluation. While CPR's primary audience is employers and other health care purchasers, the Framework could also be of use to health plans and providers who implement their own programs, as well as to benefit consultants and others asked to evaluate programs on behalf of health care purchasers.

In some cases, purchasers may wish to use this Framework to evaluate programs they have initiated directly with health care providers. In other cases, purchasers may ask a health plan to respond to the questions in the Framework as part of a process for evaluating whether the purchaser wants to partner with the plan on just one program (in the case of an existing relationship with the plan) or for all services offered (as a prospective customer). In addition, purchasers may want to use responses to the questions in the Framework to help them understand in advance the administrative and financial implications, should they decide to enroll their population in a program.

Health plans may also find this evaluation framework helpful. Not only could they use it as a starting place for a consistent approach within their own companies, or across their industry, but they can also refer to it as a guide to what employers and other health care purchasers want them to track and report.

Some health care providers may be in a position of implementing payment reforms within their own organizations. For example, an accountable care organization would need to determine how to pay its participating providers to pass on the incentives it experiences through a shared savings or shared arrangement. Such an ACO may wish to track the impact of its payment approach and use this evaluation framework as a starting point.

Furthermore, benefit consulting firms are asked regularly by their employer-purchaser customers for help in in understanding the success of various health care delivery and payment reform programs. In turn, benefit consultants query health plans and health care providers regularly about their results. They could choose to use this framework as a common ground and starting point for evaluations, thereby streamlining the administrative process and burden for plans.

We hope this tool will be useful to many parties and will serve to establish consistency in how various stakeholders evaluate their payment reform programs, allowing for more robust comparisons across programs and reducing the unnecessary reporting burden on program sponsors.

When is it Appropriate to Use the Framework?

CPR's Payment Reform Evaluation Framework can be used for various purposes and at different times in the life of a payment reform program. For instance, it can be used at the start of a new program to design the program's evaluation, or it could be used to determine the impact of a well-established program from its inception to the current period. Depending on the scope, scale, age, and intensity of the program, the outcomes portion of the Framework could be used as frequently as quarterly and as infrequently as annually.

Similarly, the Framework could be used in each of the variety of ways an employer or other health care purchaser can participate in a payment reform program:

Health Plan-Developed Payment Reform Program

Purchasers can participate in an existing payment reform program offered by their health plan(s). A purchaser may elect to gain access to such program "as-is," which may limit the ability to specify the terms of the evaluation.

Jointly Developed Program between the Health Plan and Purchaser

A purchaser may also work with a health plan to leverage its existing provider network and contracts and collaborate to establish specific payment reforms and evaluation arrangements. Part of the program development must include the design of the evaluation. Determining the evaluation specifications up front means the required data will be collected and there is agreement among parties to share such data.

Purchaser- Health Care Provider Direct Contracting

Purchasers may contract directly with a physician group, hospital or other facility or integrated delivery system to implement a payment reform program. In this scenario, the purchaser would bypass its health plan to contract directly for services from the health care provider, even though the purchaser will rely on its health plan or another third-party administrator to process claims and provide related services. The contract could encompass all health care services or just those associated with a specific condition or event (e.g. a specific chronic illness, or patients needing a particular procedure). The purchaser will need to consider what contractual terms must be in place to conduct an adequate evaluation.

Help CPR Improve the Payment Reform Evaluation Framework Over Time and Learn from Evaluations

CPR's Payment Reform Evaluation Framework reflects the thinking of some of our nation's leading experts in payment reform and program evaluation. However, there is no single right way to evaluate a health care delivery or payment reform program. Additionally, how we evaluate programs may need to change over time as new payment and delivery reforms emerge. We hope this Framework advances the consistency and rigor with which the health care industry is evaluating its efforts to reduce costs and improve the quality of care. If you have ideas for how we can improve the Framework, please give us feedback.

Furthermore, CPR hopes that the Framework can help create a more intense learning environment that enables the health care industry to spread successful approaches more rapidly and avoid design flaws and other pitfalls that evaluations may uncover. In addition for sharing ideas for how to improve the Framework, we also invite you to share the results of your payment reform programs by submitting entries into CPR's National Compendium on Payment Reform.