



DECEMBER 2015

# Developing a State-based Quality Measurement Program Using an Episode-of-Care Framework: *Recommendations for State Purchasers*

Prepared by Francois de Brantes, MS, MBA, Health Care Incentives Improvement Institute (HCI<sup>3</sup>)

## Introduction

As the US health care system moves towards value-based payment, it becomes clearer that, while alternative payment models are important, the underlying information processes required to vivify these new payment models are equally critical to the success of the payment model. As much as Patient Centered Medical Homes, Accountable Care Organizations and episode-based payments matter conceptually, the real effort lies in reforming the nature of health care information, or these payment models will languish. Significant gaps in quality of care measurement continue, as do the means for capturing quality of care data and marrying them to cost of care data.<sup>1</sup>

As a system designed for fostering accountability, federal value-based purchasing (VBP) programs have focused on the clinical outcomes of care that rely on Medicare's Physician Quality Reporting System (PQRS)<sup>2</sup> and Hospital Inpatient Quality Reporting (IQR) System<sup>3</sup>, and in some instances, in concentrated local pilots. Both the PQRS and Hospital IQR systems are conveyed through different conduits as defined measures of care. CMS integrates and reports the data in comparative data sets on physician and hospital performance respectively, largely focused on measures of care for Medicare patients.

**Whether or not genuine transformation of the delivery system takes place through the use of new payment models will depend almost entirely on the ability of practicing physicians to have access to timely, reliable and actionable feedback loops on clinical and financial outcomes.** One area where this appears to be paying off in the Medicare program is the penalty for excessive hospital readmissions.<sup>4</sup> By aligning penalties for excessive readmissions with specific comparative reports on hospital performance, CMS has seen reductions in excessive admissions for Medicare patients. State-led efforts can take a cue from this success: incentives coupled to actionable feedback reporting have the potential to give frontline clinicians the tools they need to redesign care.

This Brief outlines action-oriented steps for state purchasers to develop a quality measurement program based on episodes-of-care that leverages existing information technology infrastructure and clinical registries. Specific suggestions for state purchasers include:

1. Start with an incremental approach to quality measurement and pilot, using manual processes.

## ABOUT STATE HEALTH AND VALUE STRATEGIES

State Health and Value Strategies, a program funded by the Robert Wood Johnson Foundation, provides technical assistance to support state efforts to enhance the value of health care by improving population health and reforming the delivery of health care services. The program is directed by Heather Howard at the Woodrow Wilson School of Public and International Affairs at Princeton University. For more informations, visit [statenetwork.org](http://statenetwork.org).

## ABOUT THE ROBERT WOOD JOHNSON FOUNDATION

For more than 40 years the Robert Wood Johnson Foundation has worked to improve the health and health care of all Americans. We are striving to build a national Culture of Health that will enable all Americans to live longer, healthier lives now and for generations to come. For more information visit [www.rwjf.org](http://www.rwjf.org). Follow the Foundation on Twitter at [www.rwjf.org/twitter](http://www.rwjf.org/twitter) or on Facebook at [www.rwjf.org/facebook](http://www.rwjf.org/facebook).

## ABOUT HEALTH CARE INCENTIVES IMPROVEMENT INSTITUTE

Health Care Incentives Improvement Institute (HCI<sup>3</sup>) is a not-for-profit company that designs and implements innovative payment and benefit plan design programs to motivate physicians, hospitals and consumer-patients to improve the quality and affordability of care.

2. Create and publish a master list of data elements required from selected quality measures to appropriately identify current data collection efforts and potential gaps in measurement.
3. Create a central database that leverages existing clinical data registries and utilizes direct provider submission.
4. Develop provider feedback loops that incorporate episode-of-care efficiency metrics, with episode-of-care outcome metrics and synthesize results in a transparent manner.

### State-led VBP: Works in Progress

For states leading the way in value-based purchasing, a “pardon our dust” sign should be considered, which is to say, a work in progress is just that. There is no need to leap to artificial or stopgap measures to give the appearance of completion. Indeed, by rushing towards badly-fashioned, readily and/or publicly available mechanisms that give the appearance of completion, states actually distort information or make it too remote and ambiguous for consumers and providers alike. States need to be frank about shortfalls in publicly reported measures and resist filling them in with measures that can lead to false positives and false negatives (classifying a hospital as being good at everything when it’s not or bad at everything when it’s not).

By emphasizing episode-of-care pathways, as the states of Arkansas, Ohio and Tennessee are doing,<sup>5</sup> gaps in quality measurement can be identified, and where need be, uniquely redesigned. States can address the gaps incrementally and make the most of limited resources by building episode-specific measure sets.

A case example for this incremental approach can be found in the work of Community Health Choice (CHC), a Houston-based Medicaid plan. CHC launched a “womb-to-crib” bundled payment program and tied all of the phases of pregnancy, delivery and newborn care into a single, severity-adjusted global fee. When the plan looked for available data on quality of maternity care, data available to CHC at the time were fragmented and limited. As a result, the plan created a maternity quality scorecard which requires input from clinical record data. Participating providers use manual processes to submit information from medical

records as an initial step. Once results are validated and found useful for clinicians, automated processes can be instituted. Ideally, over time manual processes such as these will get converted to automated data feeds using clinical registries<sup>6</sup> as discussed later in this Brief.

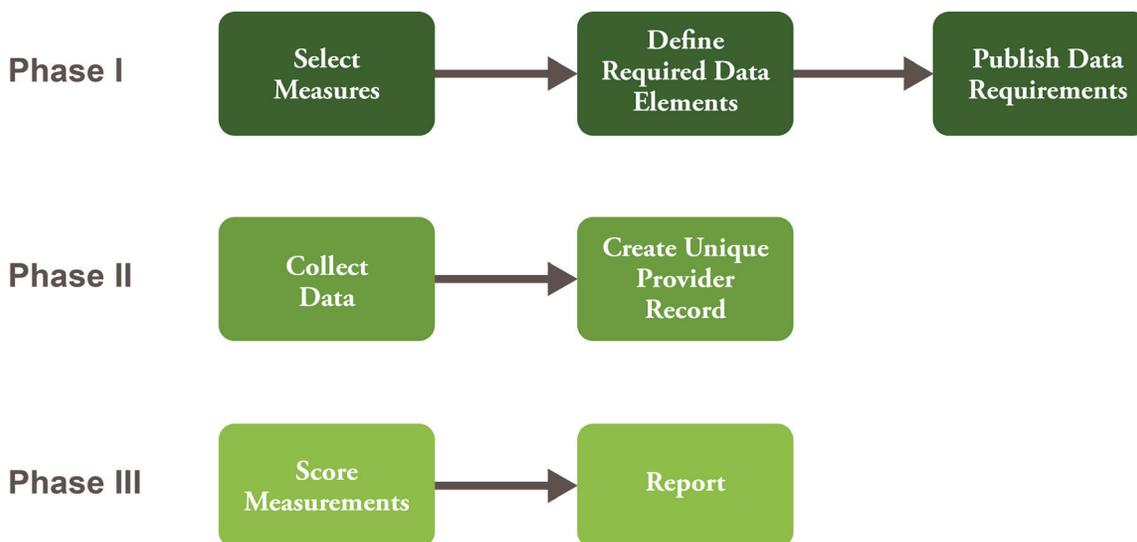
Designing a quality scorecard that matches the scope of the bundle is an essential feedback mechanism for clinicians, providing two critical views of the same episode of care: a financial view and a quality of care view. These views are within the clinicians’ line of sight and highly actionable, making care redesign and other process changes far more likely to happen faster.

Relying on manual processes to get started and ensure provider engagement, when registries are not available, is defensible to engage providers on quality performance. A manual process allows for refinement and modification, and requires only minimal capital investment and modest amounts of labor. Once methods are proven, scaling issues become important, but not until then.

Given the dearth of publicly available measures on the quality of most episodes of medical care,<sup>7</sup> states must roll up their sleeves, work with local provider advisory groups, and develop ad hoc protocols for data collection and reporting. While few meaningful measures are publicly reported, quality measures have been defined for a large number of medical episodes of care and a reasonable subset are being reported and collected through clinical data registries.<sup>8</sup> As part of the technical assistance for Tennessee Medicaid,<sup>9</sup> HCI<sup>3</sup> delineated the availability of measures and the corresponding registries collecting them related to episodes targeted by the state. **Appendix A** provides examples of clinical data registries (CDRs), including those qualified by Medicare, which align with certain episodes of medical care.

### Process for Quality Measuring and Reporting

The following section outlines a three-phase pathway (Figure 1) for establishing, measuring and reporting protocols that enable states to create timely clinical outcome feedback loops by leveraging existing data registries and providing alternative data submission protocols for providers who do not have access to or choose not to use available registries.

**Figure 1: Potential Data Sources and Approach for Quality Reporting**

### ***Phase 1: Selecting Performance Measures and Defining Data Elements***

As noted in *Figure 1*, the first step involves selecting quality performance measures. Using standardized measures and common measure sets reduces the administrative burden and sends a consistent message about provider performance accountability. For additional perspectives on selecting measures, states may wish to review a prior SHVS brief, *“Considerations for State Development of Performance Measure Sets.”*<sup>10</sup>

*Create and publish a master index of candidate data elements:* States should examine clinically related or proximate episodes to reduce potential duplication of data elements being measured. The process for developing performance measures begins with a) the element being measured, for instance, systolic blood pressure, and b) the patients that should be included (and excluded). Data elements for measure sets of related conditions may be used for multiple measures. For example, a measure set often includes measures of superior control (such as number of patients with systolic blood pressure below 120) and measures of poor control (such as number of patients with systolic blood pressure over 140). Noticeably, both of these examples are measuring the same clinical indicator: systolic blood pressure, which can then be used to create a number of quality measures across many episodes of medical care. **It’s essential to create a master index of candidate data elements to determine the overall quantity of such data elements and better indicate to physicians and hospitals the extent of the data collection process.** Publishing a master index helps

all involved with a state effort to determine which data elements they are currently collecting and to identify potential gaps. Gaps can be assessed both in terms of the extent to which those providers for whom the measures will be applied are collecting the data elements, and the number of data elements that need to be collected to create all agreed-upon measures. The result should enable stakeholders to prioritize data collection efforts.

Publishing the list of desired data elements also signals clinical data registries and Electronic Medical Record (EMR) vendors of upcoming demands from physicians and hospitals on extracting data from internal medical record databases. For the vast majority of existing clinical quality measures, required data elements reside, in some fashion, in existing and deployed EMR systems.<sup>11</sup> Our experience suggests that extracting needed data elements from practices, hospitals and health systems with an EMR is not a particularly big challenge. The key is to be clear on the data elements and any other specifications related to a measure for which the data element will apply, for example clinical exclusions. Alternatives to EMRs are discussed in the next section.

### ***Phase 2: Data Collection for Quality Reporting***

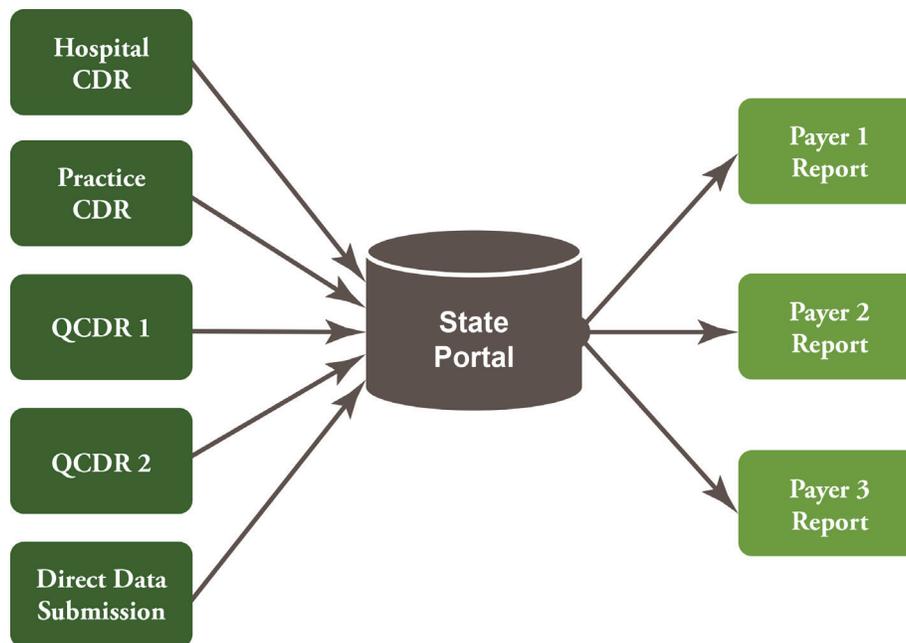
**Whether measurement data comes from established registries, directly from providers, or participating health plans, it should be subsumed into a master database and reconciled around single provider records.** Assembled data can then be analyzed to compare the effectiveness of treatments and reported out to providers in a consistent way, irrespective of payers to the extent feasible. This concept is

important because the traditional way in which provider performance measurement has been conducted is payer by payer. As a result, provider performance reporting has a tendency to vary by payer, creating confusion.

A centralized scoring mechanism across all of a provider's patients will ensure that feedback to the provider on the quality of care will be the same across all payers. As part

of technical assistance for Tennessee, HCI<sup>3</sup> suggested a data collection and reporting schema as depicted in Figure 2, where the inputs come from hospital and practice Clinical Data Registries (CDR), CMS-authorized CDRs (known as Qualified Clinical Data Registries (QCDR)), and/or direct data submissions from providers, and the outputs are reports to clinicians.

**Figure 2: Potential Data Sources and Approach for Quality Reporting**



QCDRs<sup>12</sup> are registries authorized by CMS to collect quality measures from physicians to satisfy reporting requirements of the Medicare Physician Quality Reporting System. As such, leveraging QCDRs can speed up the process of setting up a data collection infrastructure. Generally speaking, leveraging registries – whether native to an EMR in a provider organization, maintained by a medical specialty society, or qualified by CMS – is the more efficient and effective way of developing a central data collection system. Direct data submission by providers requires instituting a series of processes, including data validation and integrity, that have to be designed from scratch.

#### **Basic decisions for states relative to designing direct data submission portals include identifying:**

- *The purposes of the portal* – Data submission only or data submission and reporting;
- *The scope of the portal* – Whether direct submission will be accepted for all measures/programs or only some;
- *Whether access by parties other than the clinician managing the patient will be allowed* – Many physicians may elect to

have a practice administrator submit data on their behalf;

- *What auditing requirements for sampling of patients included in the direct data submission are necessary* – Typically, direct data submission entails drawing the data from a random sample of patients rather than reporting on all patients.

#### ***Phase III: Measure Scoring and Reporting***

No matter their good intentions, states getting into the process of scoring and reporting on performance should be aware that the physician community tends to view publicly reported clinical and financial performance with deep suspicion. In addition, two decades of measurement reporting have shown that those being measured gravitate towards emphasizing measures that are common with easily attainable thresholds. This has been true at both the federal and state level. Today, little usable physician and hospital quality information exists for the public at large.<sup>13</sup> **As a result, state purchasers should keep these important lessons regarding performance measurement and reporting in mind:**

1. *Measure what matters* – Scorecards should be concise and populated with high impact measures that have a direct relationship with patient outcomes.
2. *Encourage continuous performance* – All measures should be scored using the result of the numerator/denominator calculation, and that result should be applied against the total number of points allocated to each measure. Additions to numerators should yield additional points, so that clinicians have continuous incentives to improve the quality of care.
3. *Make results actionable* – Feedback should be timely and relevant. This means: (a) providing benchmark comparisons and best practice sharing; (b) making clinical reengineering experts available to frontline clinicians; (c) providing knowledge exchange mechanisms to facilitate peer-to-peer interactions (such as online forums).
4. *Make results and reports consistent* – Whenever feasible, states should assess quality of care across payers, not payer-by-payer. Assessing provider performance across all patients avoids a potential sample selection bias and the likelihood that a physician will have varying scores from one payer to another.

*Integrating, Not Reconciling Data Streams:* State agencies spearheading these efforts should be cognizant of the fact that there is a good chance discrepancies will appear between the clinical exclusions/inclusions of defined quality measures and the corresponding episodes of care definitions. For example, patients who have undertaken two-step therapies for controlling their blood pressure and who still have high blood pressure can, under certain circumstances, be excluded from a quality measure. However those patients will always be included in an episode of care for several reasons. First, by default, because there is no way to discern such an exclusion from claims data, and second because the quality measure is designed to measure the effectiveness of the physician's treatment of the patient's condition, while episodes of care cost accounting is designed to measure the efficiency with which a physician manages patients with a certain condition. For the latter exercise there is no rationale to exclude patients who have taken two therapies and can't get their blood pressure under control. The payer still has to pay for the costs of care.

### Sustainable Feedback Loops: The Real Goal

Over the past decade or so, the use of the term “feedback loop” has increasingly entered health policy. A feedback loop from a quality measurement perspective is a way in which physicians can understand their performance, relative to a

benchmark. The underlying assumption of a quality measurement program is that the physician would change behavior to improve their own performance based on the feedback. In **Appendix B**, we outline necessary system parameters common to viable feedback loops that states should keep in mind when designing quality reporting mechanisms.

**Insofar as transparency is concerned, state purchasers should set up a performance reporting system that synthesizes cost (efficiency) and quality (effectiveness) in a way that concisely reveals value to payers, providers and consumers.** In developing a transparency approach, states should recognize that each of these stakeholders has different interests and levels of understanding. The value synthesis rests on combining efficiency calculations (total episode cost against benchmarks) and effectiveness calculations (episode-specific patient quality of care against benchmarks), and feeding back the resulting value synthesis to all providers and other stakeholders.

### Claims and Clinical Data

State purchasers can think of data drawn from claims data as Channel 1 (measuring efficiency), and non-claims, clinical data as Channel 2 (measuring effectiveness). Units of analysis for Channel 1 are patient-centered episodes of care, with an eye towards measuring variability in these episodes. Episode cost variability can come from several sources: the price of individual services, the use of services (either too many or too few), and the mix of services. Information on the contribution of each of these sources to the total variability in episode costs can help providers better understand how to improve the sum of the inputs used to manage an episode of medical care. The importance here is not simply in creating a feedback loop on a provider's specific variability, but rather how that variability compares to that of others. For example, a provider who gets a report that shows her variability comes mostly from higher pricing of services will have a very different strategy than a provider getting a report indicating that his variability comes from a significantly higher use of certain services. As one might surmise, these reports should be payer specific, especially when analyzing variability based on price.

The units of analysis for Channel 2 (non-claims, clinical data) are all patients that have a specific medical episode, irrespective of the payer, and for two principal reasons. First, states should want to encourage providers to treat all patients with a certain condition as optimally as possible and not introduce a potential payer-specific bias. The central idea being that a single provider quality score cannot be manipulated by a payer to try and tilt that provider's

attention preferentially towards that payer’s plan members. Second, insofar as transparency is concerned, states should set up a system that synthesizes cost and quality in a way that succinctly reveals value to payers, providers and consumers, with each having different interests and levels of understanding. **The value synthesis rests on combining efficiency calculations (total episode cost against benchmarks) and effectiveness calculations (episode-specific patient quality of care against benchmarks), and feeding back the resulting value synthesis to all providers.**

## Conclusion

While the concept of tying cost and quality of care into a timely, actionable and reliable report for physicians seems common sense enough, the general availability of data to create these reports is extremely low. As such, states that wish to accelerate the transformation of the existing delivery system into one that delivers high quality and affordable health care have to take action to develop a comprehensive

data collection and reporting mechanism.

This Brief suggests that such an approach be done using episodes of medical care – such as a chronic condition, an illness or a major treatment/procedure – as the central unit of measure because (a) quality measures are generally tied to specific episodes of medical care, and (b) acting on the cost of an episode of care is a lot easier to do for frontline clinicians than acting on a higher level of cost aggregation such as total cost of care. Of course, for states implementing bundled payment programs, the cost of the medical episode is simply the price of the bundled payment.

Further, this Brief outlines specific steps that can be taken by states to launch a data collection and reporting effort, perhaps with manual processes initially, and then to scale such an effort. The information technology infrastructure in place in the US today can be leveraged to rapidly scale a central data collection and reporting process and create highly relevant feedback loops for providers.

## Appendix A: Sample of Select Episodes of Care and Related Clinical Data Registries

| Episode                     | Matching Qualified CDRs (QCDRs)*  | Matching CDRs  |
|-----------------------------|---|--|
| Asthma acute exacerbation   | American Academy of Allergy Asthma and Immunology (AAAAI)<br><a href="https://www.aaaai.org/home.aspx">https://www.aaaai.org/home.aspx</a>  |  |
| Bariatric surgery           | Metabolic and Bariatric Surgery Accreditation and Quality Improvement Program   |  |
| Breast cancer               | American Society of Breast Surgeons Mastery of Breast Surgery Program<br>American College of Physicians (ACP) Genesis Registry<br><a href="https://www.medconcert.com/content/medconcert/Genesis/">https://www.medconcert.com/content/medconcert/Genesis/</a> |  |
| Depression                  | ACP Genesis   |  |
| Diabetes acute exacerbation | ACP Genesis<br>Chronic Disease Registry   |  |
| Female reproductive cancer  | American Society of Clinical Oncology (ASCO)<br>QOPI<br>Oncology Nursing Society Quality Improvement Registry<br>Oncology Quality Improvement Collaborative   |  |
| Lung cancer                 | ASCO QOPI<br>Oncology Nursing Society Quality Improvement Registry<br>Oncology Quality Improvement Collaborative  |  |
| Neonatal                    |   | Vermont Oxford VLBW Database<br><a href="https://public.vtoxford.org">https://public.vtoxford.org</a>                                  |
| Perinatal                   |   | American Association of Birth Centers (AABC) Perinatal Data Registry<br><a href="http://www.birthcenters.org">www.birthcenters.org</a> |
| Spinal fusion               | Anesthesia Quality Institute: National Anesthesia Clinical Outcomes Registry  |  |
| Tonsillectomy               | Anesthesia Quality Institute: National Anesthesia Clinical Outcomes Registry  |  |
| Total joint replacement     | American Joint Replacement Registry   |  |

\* **Qualified CDRs:** clinical data registries authorized by CMS to collect quality measures from physicians to satisfy reporting requirements of the Medicare Physician Quality Reporting System. For a list of 2015 QCDRs see: <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/PQRS/Downloads/2015QCDRPosting.pdf>

## APPENDIX B: 7 Essential Questions That Identify Working Feedback Loops in Healthcare

1. *Where are the circuits of data and information connecting providers?* – Given the fragmented payer and provider institutional arrangements, siloed information systems, and inconsistent means of data collection, it is hard to discern the structured conduits connecting the relevant healthcare actors. The systemic “wires” must be in place.
2. *How is work (output) measured?* – This question would fall into two parts: a) the definition of work, and b) the salient contributors to work. There are so many parties, both governmental and private, creating inconsistent quality measures, the result can only be confusion and lack of uptake. Moreover, the two salient contributors to work, patients and providers, are treated as if they live on different planets. Almost all the measures place heavy emphasis on provider response, with little attention to patient response. In payment reform, this asymmetry begs for correction.
3. *What is the unit of analysis?* – Feedback systems operate on meaningful units of analysis, and thus the unit of analysis has great bearing on work measures. If the work measures are analyzed through inappropriate units, as with hospital-only measures, analysis and work output fall out of sync with each other. The appropriate unit of analysis in healthcare cannot, therefore, be institutional; it has to focus on the primary consumer of the work product: the patient.
4. *How much energy is consumed?* – Engineers are in a constant quest to lower the amount of energy required per unit of work; this is the definition of efficiency, and is often quantified in terms of wasted energy. Systems engineers would be staggered by how much energy is wasted in American healthcare, the current of work being dollars. Dollars, therefore, tie work measured and unit of analysis together as definable work products. FFS and TCOC are not defined healthcare products in dollar terms if the patient is the unit of analysis.
5. *Are the feedback mechanisms parsimonious?* – Not all metrics are equal. At some point, measuring every conceivable variance to the  $n^{\text{th}}$  degree and granting them equal weight creates more noise than signal. It turns out that most episodes of care have only a handful of meaningful metrics, that when controlled for, give the most amount of bang for the buck. This is what is meant by creating high signal to noise feedback loops. A parsimonious design gives relevant decision-makers the right amount of data points (signal) they need to optimize outcomes (work product), and weeds out extraneous information (noise).
6. *Is the feedback timely?* – This system parameter seems fairly obvious, in that outdated feedback is not only useless; it’s a nuisance. Actionable feedback must not only be parsimonious, it must be available at critical decision nodes where applying it has the most amount of potential to affect optimal Delta.
7. *Where are the control mechanisms?* – The means of making operational adjustments to bring actual performance to optimal performance (Delta) are either nowhere to be found (as with FFS), or posited in structures so large and ill defined (as with ACOs), as to conclude there are no controlling mechanisms, at least none that could qualify as actionable feedback systems. And this brings us to the heart of the matter: accountability. Since we’re not talking about feedback in machines, but rather, feedback within human networks and relationships, then accountability must be aligned with control, and that means getting the first 6 parameters right; otherwise, managerial spans of control, or “lines of sight,” become diffuse, chaotic and very difficult to coordinate.

## Endnotes

---

- <sup>1</sup> Report from the National Quality Forum: 2012 Measure Gap Analysis; [www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=72981](http://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=72981).
- <sup>2</sup> <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/PQRS/index.html?redirect=/pqri/>.
- <sup>3</sup> <https://www.cms.gov/medicare/quality-initiatives-patient-assessment-instruments/hospitalqualityinits/hospitalrhqdapu.html>.
- <sup>4</sup> See for example “Transitional Care Interventions Prevent Hospital Readmissions For Adults With Chronic Illnesses”, Kim J. Verhaegh et al. *Health Affairs* September 2014 vol. 33 no. 9 1531-1539.
- <sup>5</sup> Arkansas, Ohio and Tennessee have launched statewide bundled payment programs for specific episodes of care as the central focus of their Medicaid payment reform efforts. For additional detail on the Tennessee initiative, see: <https://www.tn.gov/hcfa/topic/episodes-of-care>.
- <sup>6</sup> Registries are databases containing specific information on patients and have been instituted by Medical Specialty Societies to help their members better monitor patient outcomes and understand the effectiveness of treatments. Some registries are also native to electronic medical records and are simply a subset of data stored in EMRs, making it easier for clinicians to extract information.
- <sup>7</sup> See HCI<sup>3</sup> report, “State Report Card on Transparency of Physician Quality Information,” December 16, 2014.” at <http://www.hci3.org/content/physician-quality-transparency-report-2014>.
- <sup>8</sup> Many clinical data registries exist and are often condition-specific. For example, the Oncology Quality Improvement Collaborative (<https://www.medconcert.com/content/medconcert/ONCQIR/>) measures and reports on outcomes in oncology and specialty care, whereas the Vermont Oxford Network hosts a database about the care and outcomes of high-risk newborn infants (<https://public.vtoxford.org/databases/very-low-birth-weight/>).
- <sup>9</sup> Support for this technical assistance work in Tennessee was provided through a grant from the Robert Wood Johnson Foundation’s State Health and Value Strategies program.
- <sup>10</sup> See <http://statenetwork.org/resource/considerations-for-state-development-of-performance-measure-sets/> accessed November 2015.
- <sup>11</sup> HCI<sup>3</sup>, through its Bridges To Excellence (BTE) programs, has been successfully collecting data elements for dozens of quality measures on common chronic conditions from various EMR systems for well over five years. For more BTE information see: [http://www.hci3.org/what\\_is\\_bte](http://www.hci3.org/what_is_bte).
- <sup>12</sup> For a list of 2015 QCDRs see: <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/PQRS/Downloads/2015QCDRPosting.pdf>.
- <sup>13</sup> See HCI<sup>3</sup> report, “State Report Card on Transparency of Physician Quality Information,” December 16, 2014.” at <http://www.hci3.org/content/physician-quality-transparency-report-2014>.