

Performance Measurement and Reporting at the Clinician Level: Improving Its Potential While Developing More Comprehensive Care Accountability

EXECUTIVE SUMMARY

At the behest of employers, many health plans have made clinician-level measurement results publicly available, part of pay for performance, and factors in health plan network tiering strategies in order to create more urgency for improvement and to meet increased consumer demands for clinical performance information. These steps have dramatically raised the stakes, drawing more scrutiny to current measurement efforts. The early results of that scrutiny have led to calls for more transparent, valid, reliable, standardized, and actionable measurement. Although holding clinicians accountable for all of a patient's care has its utility and areas of high validity, many instances, or episodes, of care for a patient involve elements that are not under the direct control of a single clinician.

Clinician-level measurement is undergoing refinement. The shortcomings of current measurements are driving concurrent efforts to more broadly define relevant, important, and measurable elements of condition-specific episodes of care and related accountability.

Preface

Millions of times each day, patients interact with the U.S. healthcare system. During these interactions, most patients receive the benefit of solid clinical judgment and technical expertise from their care providers and witness basic and state-of-the-art technology appropriately applied. However, some patients do not receive these benefits and instead experience care that reflects the system's weaknesses, including the provision of unnecessary or ineffective services and the lack of provision of needed services. These weaknesses might be eliminated through measurement and the production of actionable information, bringing each patient one step closer to optimal care.

Managing clinicians play a dual role within a healthcare system, serving both as suppliers of services to patients and as decisionmakers as to the use of healthcare resources. Current incentives encourage *more* care rather than the *right* amount of care. In other industries, reducing waste often benefits the bottom line. By contrast, in much of our fee-for-service healthcare

➔ *Continued on page 2*

system, reducing waste threatens personal provider income and business entity margins. The dominant care provider business models seek to maximize revenue and income. Examples of this phenomenon at the clinician practice and group level include the acquisition and deployment of new revenue-generating technology (e.g., extremity MRI machines in primary care offices) and clinician ownership of ambulatory surgical centers. In the latter case, a recent study has found that “physicians at physician-owned facilities are more likely than other physicians to refer well-insured patients to their facilities and route Medicaid patients to hospital outpatient clinics.”¹

With knowledge of significant clinician practice level quality variation and seeking to increase the provision of optimal care, many of healthcare’s stakeholders, including clinicians and consumers/patients and their organizations, have supported or deployed improvement strategies, including clinical measurement, transparency of measurement results, differential pay for differential performance, and the shifting of increasing numbers of patients to higher-performing clinicians. These and other, similar activities have led to performance incentives for physicians and other clinicians who manage care. Although measurement at the clinician level may be appropriate for many performance measures, holding individual managing clinicians exclusively accountable for all elements of care is seen as an oversimplification, given the complex dynamics of the healthcare system and the multitude of care providers who are frequently involved in a patient’s care. With notable exceptions, broadly accountable care entities, capable of accepting overall responsibility for patient care, are as yet only partially invented and narrowly deployed.

Meanwhile, today’s clinical performance measurement can be alternately described as bold yet controversial, as helpful, yet at times misleading, and also as conceptually transformative but constrained by poor alignment of economic incentives. As a result, near-term exigencies exist to

improve upon what has been developed and deployed to date, while addressing the pressing need for more comprehensive frameworks of accountability.

Background on Current Efforts

Purchasers, payers, and consumers of care are urgently seeking relevant and reliable measurement of clinicians’ practices and useful depictions of practice performance to guide their choices of care providers. For many of these stakeholders, the goals of such measurements have been threefold: to improve process quality, outcomes, and resource use.

Process quality measurement is best illustrated historically by the Healthcare Effectiveness Data and Information Set (HEDIS)² promulgated by the National Committee for Quality Assurance.³ Originally referred to as the Health Plan Employer Data and Information Set, HEDIS was created in the 1990s in response to employers’ calls for a standard means for assessing health plans. Historically, HEDIS was focused on the performance of health plans with respect to preventive and primary care delivery, such as immunization rates and performance of breast cancer screening by mammography. The performance of health plans on HEDIS measures has long been a key element in health plan accreditation.

More recently, physicians, physician specialty societies, and physician-led performance improvement⁴ initiatives such as the Physician Consortium for Performance Improvement have responded to the call for additional measures to expand the conditions and specialty breadth of measurement. These efforts have contributed to the current 210 clinician-level measures currently endorsed by the National Quality Forum (NQF).

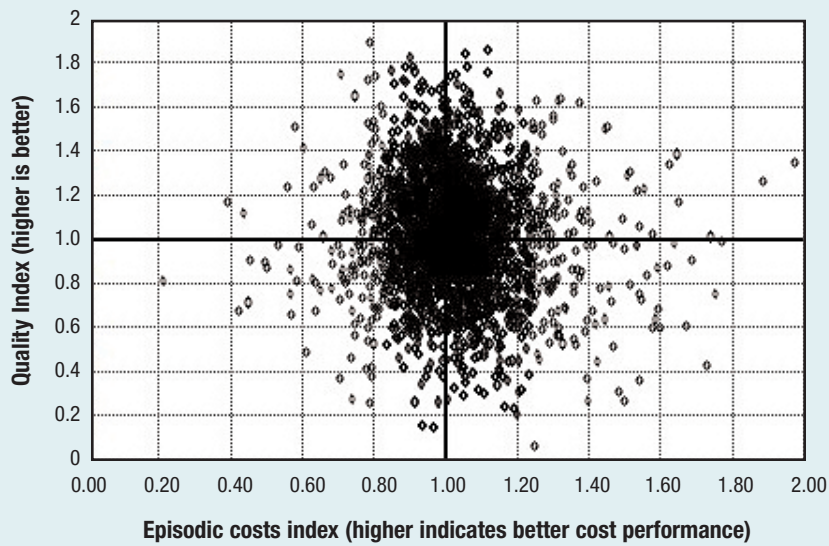
Measurement of clinician-level resource use, most commonly depicted as costs, has also been a component of health plans’ clinician report cards. However, early measurement of resource use was typically focused on units of services, such as the number and type of office visits, the

number and type of hospital admissions, and the use of lab, x-ray, and pharmacy services across the clinician’s practice.

In the last few years, measurement has shifted to a patient- and condition-centric approach based on episodes of care. Currently, episodes of care are created by proprietary software algorithms⁵ that evaluate claims submitted by care providers for payment. These software tools aggregate resource usage for the evaluation and treatment of an individual patient’s specific condition. Software tools use different approaches to define episodes. For example, some tools use an arbitrary time period (e.g., 12 months from the onset of symptoms), while others analyze claims data to identify patient-specific end points (e.g., three months with no healthcare services). The resulting episode cost is one way to express resource use measurement and has been used by payers to compare clinician performance. Assigning episodes of care to individual physicians remains problematic when more than one managing clinician is involved in an episode of care. Recognition of this difficulty has further advanced the call to define more inclusive and accountable entities that aggregate clinicians engaged in the same patient’s care.

Initially, large employers were hopeful that process quality would correlate with cost-efficiency, as was often found to be the case within in their own industries. However, in 2004, studies began to emerge that demonstrated that there was no correlation between the performance of clinicians with respect to process quality and their performance with respect to cost-efficiency.^{6,7} These early studies identified variation in measured care performance using commonly deployed tools at the clinician practice level and at the medical group level in aggregate. Four-quadrant graphs depicting process quality and episode of care cost variation (see Figure 1) have been powerful graphics for employers seeking to move business to higher performers. In this example, each point on the graph depicts a Washington State clinician as compared to his or her peers. The study underlying this data used common quality

FIGURE 1 Clinicians' Quality and Episodic Costs



Each data point represents a Washington State clinician's practice results for process quality measures and episode of care-based cost-efficiency.

Source: The Regence Group, used with permission, all rights reserved, 2004.

measures and episode treatment graphs to generate expected/observed indices of process quality and episode-based costs of care.⁶

Efforts are now under way to supplement the process quality and cost information that is available through claims data, with additional process and patient outcomes information available from electronic health records, registries, and patient surveys. Prescription records and laboratory results augment medical claims and provide additional clinical depth to measurement. Certain clinical process measures and outcomes may be abstracted from the clinical record in support of specialty-specific registry efforts (such as those of the Society of Thoracic Surgeons) or governmental mandates. Health plans are increasingly leveraging clinical process measures and outcomes from such registries in their efforts to assess clinician performance.

Clinician-level measurement of outcomes remains challenging due to inadequate capture of necessary clinical data, and because intermediate to long-term outcomes are affected to greater degree by

non-clinician factors, including patient characteristics. One area of clinician measurement, patients' experiences with care, does appear to be ready for deployment as a care outcome measure. Safran et al. demonstrated the feasibility of obtaining highly reliable measures of patients' experiences with individual physicians and practices while adjusting for patient factors.⁸

Near-Term Priorities: Refining and Improving Current Clinician Measurement Efforts

As measurement at the clinician level has grown, concerns have been raised about the fairness, validity, and reliability of such measurement systems. In the deployment of some measurement systems, inferences made about performance at the clinician level have been characterized as at times overreaching and as beyond the sophistication and capabilities of some of the tools and methods.

A collaborative effort between consumers and health professionals recently

established criteria for clinician-level measurement efforts. The *Patient Charter for Physician Performance Measurement, Reporting and Tiering Programs*⁹ (Charter) offers these criteria and details further attributes that are desirable for clinician level measurement:

1. Measures should be meaningful to consumers and reflect a diverse array of physician clinical activities.
2. Those being measured should be actively involved.
3. Measures and methodology should be transparent and valid.
4. Measures should be based on national standards to the greatest extent possible.

The Charter further calls for third-party evaluation of measurement efforts to assess compliance with the Charter principles. Notably, other stakeholder groups, including business/purchaser coalitions, have subsequently endorsed these principles.

Recently, NQF commissioned a discussion paper on clinician level measurement.¹⁰ The report was followed by a multistakeholder meeting on the subject. Several recommendations resulted from the meeting, including the following:

1. **Enrich clinical data.** Accelerate efforts to capture relevant clinical data to augment administrative claims data.
2. **Improve the breadth and relevance of performance measurement.** Increase the number and quality of measures in order to better reflect the breadth and focus of a clinician's practice. Broaden measured specialty types.
3. **Aggregate data from public and private insurers.** Increase the number of observations upon which clinician performance might be more adequately assessed.
4. **Recognize limitations; use measurement and reporting responsibly.** Measurement methods may be imperfect and introduce the "noise" of measurement imprecision. In many instances, actionable improvement steps can be identified in the presence of such imprecision. Public transparency,

provider tiering, and pay for performance inherently place higher standards on information precision. A collaborative balance with a focus on driving near-term improvement is needed given that the perfection of clinical performance measurement will be a long-term aspiration.

5. **Provide due process and continuously improve measurement and reporting.** Address the concerns of clinicians who identify potential undue methodological impact or incomplete or inaccurate data, and include the right to request review and reconsideration and provide additional data.
6. **Increase transparency and standardization of measurement and reporting.**

Both those conducting the measurement and those being measured are challenged by the presence of non-standardized data and the varying approaches that are taken to clinical performance measurement. Areas that are in need of collaborative attention and action in the near term—necessarily open to improvement and potential compromise—include the following:

- methods of identifying unique clinicians, including the challenges of identifying the specific clinicians involved in providing care;
- methods of determining physician specialty peer comparison groups;
- methods of assigning a clinician to a peer group;
- methods of attributing measures to physicians—this includes the challenges of apportioning and assigning accountability and the difficulty in attributing responsibility for medical decisionmaking when there are multiple care providers; and
- number of observations (patient-level measures assigned to a clinician) and the resulting degree of statistical soundness of performance reporting and rating systems.

Longer-Term Priorities

Clinicians and the decisions that they make have a strong influence on the delivery of optimal care. However, recent examinations of care delivery have found substantial variation in geographic and care system performance. Clinicians may have obtuse or poorly defined responsibilities for important attributes of care variation. Addressing such system-level variation has required new conceptual thinking along with the design and development of new measurement approaches and tools. The system-level variation found to date has brought further critical examination of the economics and current incentives that are powering a wasteful system of high cost and disappointing quality.

Longer-term priorities include the following:

Spanning Silos of Care with Extended Episodes of Care

Extending the concept of episodes of care to be more inclusive of other measurable domains of optimal care is seen as one way to span the silos of care, measurement, and accountability. NQF's Establishing Priorities, Goals and a Measurement Framework for Assessing Value Across Episodes of Care project is pursuing this goal.¹¹

NQF's National Priorities Steering Committee on Efficiency and Episodes of Care has completed its first draft report, entitled *Measurement Framework: Evaluating Efficiency Across Episodes of Care*.¹² Presenting a "measurement framework for evaluating efficiency across extended episodes of care," the report is intended to guide ongoing efforts to create de novo conceptual and analytical models of episodes of care that more broadly incorporate measurable elements of optimal care. Efficiency of care in this context "means a measure of the relationship of the cost of care associated with a specific level of performance measured with respect to the other five IOM [Institute of Medicine] aims of quality [safety, timeliness, effectiveness, equity, and patient-centeredness]."¹³

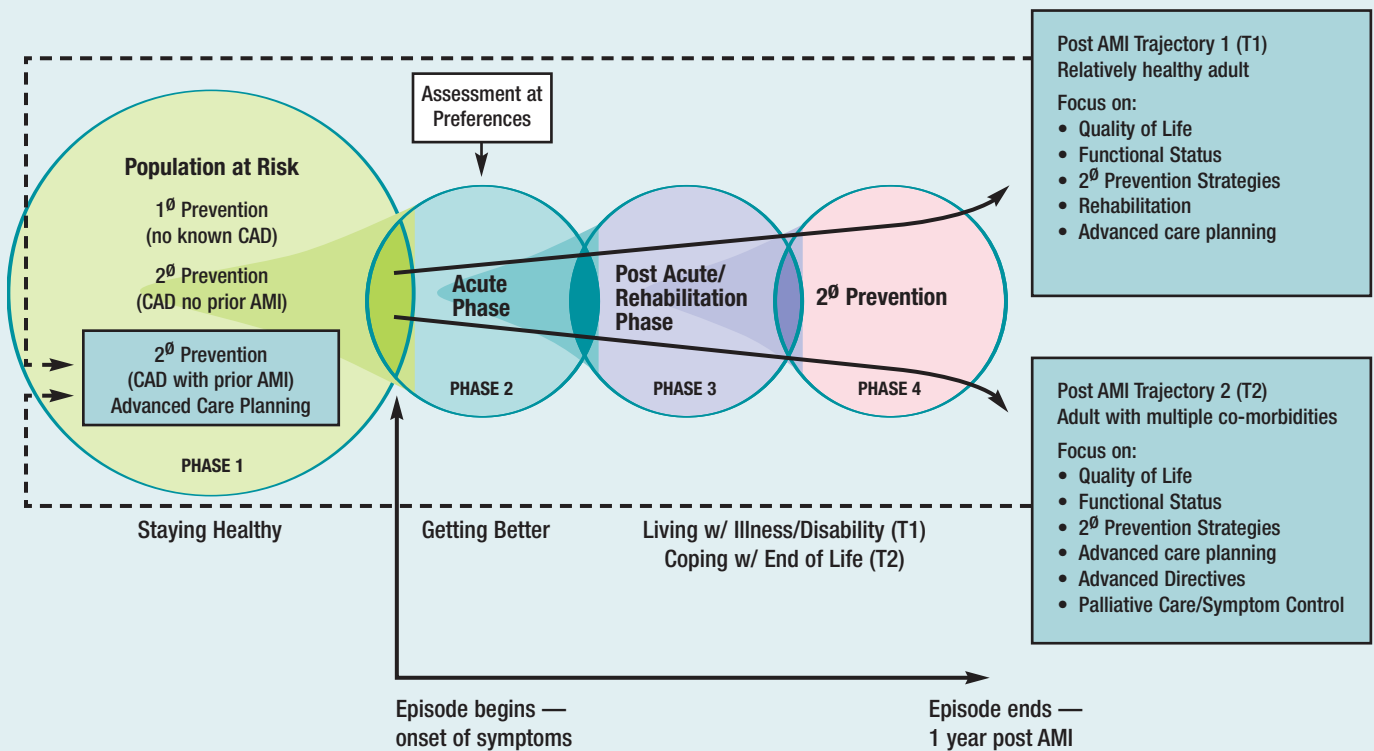
These efforts also seek to develop new public domain tools for the assessment of episode-based performance. Acute myocardial infarction served as a framework example (see Figure 2). Through the efforts of the Quality Alliance Steering Committee's High Value Healthcare Project—funded by the Robert Wood Johnson Foundation and coordinated by the Brookings Institution, and work funded by a grant to the American Board of Medical Specialties—multiple work groups are convening to create prototype extended episodes of care for common conditions.¹⁴

There is much interest in aligning payment with patient-focused episodes. Imagine a situation in which the U.S. healthcare system and its components compete on the *value of care*.¹⁵ As depicted by the AQA Alliance, the value of care "is a measurement construct of specified stakeholder's (such as an individual patient's, consumer organization's, payer's, provider's, government's, or society's) preference-weighted assessment of a particular combination of quality and cost of care performance."¹⁶ Further imagine that those preference-weighted assessments are embedded into extended episode of care analytics. Well-characterized and collaboratively built episodes of care have the potential to enable new value-based payment models at the condition-specific level.

Better Defining and Leveraging Knowledge of Team and System Impacts

One of the stronger arguments against a one-size-fits-all clinician level measurement system is the degree to which patient care for many conditions is delivered by teams of clinicians and impacted by the processes and infrastructure of care external to clinicians' practices. Teams are often ill-defined and poorly integrated, with little communication and poor coordination of care. In other instances, the team approach has been finely honed, with both individual and shared accountability among team members leading to optimal care results.

FIGURE 1 Context for Considering an Acute Myocardial Infarction Episode



At times, care teams for certain patients and conditions might be fully contained within a single clinician practice, a multiclinician medical group and its staff, a physician-hospital organization, or another integrated delivery system. Some integrated delivery systems have attained predictably consistent high levels of comparative performance in areas of focus. Such organizations embrace performance-differentiated medicine and have gone to the extent of specifying and warranting performance levels.¹⁷ However, in many locations, care teams are likely to span multiple organizations with diverse agendas and business models, differing information systems, and variable commitments to continuous quality improvement (CQI). Therein lies one of the biggest challenges to the optimization of care: the lack of transorganizational alignment around and accountability for the entirety of each patient's care.

This frequent lack of a more encompassing accountable entity is one reason why the performance of system constituents—clinicians, clinician groups, and hospitals—receives so much attention. *The challenge is to span across silos of care, measurement, and accountability when optimization of care requires it.*

Conclusion

Until we can more broadly define overall system and shared accountability with tools such as extended episodes of care, the current tools depicting performance of components of our care system, such as clinical care providers, will likely see continued and, although hopefully, refined deployment and use. Opportunities to improve quality and reduce waste exist at multiple levels within the current health-care system. Further identification of these opportunities, greater recognition of

the role of care teams, clearer paths to accountability, and further installation of CQI processes within clinical care are necessary components in the evolution of performance-differentiated medicine.

Acknowledgments

This Issue Brief is based in part on a background paper commissioned by NQF and written by Mark Rattray, MD, and also on the proceedings of a workshop convened by NQF in March 2008. Dr. Rattray wrote this Issue Brief. Sara Maddox and Tamara Lee, consultants to NQF, and Sarah Callahan, Dwight McNeill, and Katharine Torrey, NQF staff members, also contributed to this Issue Brief. ●

NQF

NATIONAL QUALITY FORUM

NQF's mission is to improve the quality of American healthcare by setting national priorities and goals for performance improvement, endorsing national consensus standards for measuring and publicly reporting on performance, and promoting the attainment of national goals through education and outreach programs.

NQF Issue Briefs provide insight into payer, policy, and industry efforts to promote quality healthcare and are produced with sustaining support from the Cardinal Health Foundation. Additional support is provided by sanofi-aventis.

This publication series is part of NQF's education and outreach initiative *Navigating Quality Forward*.

NQF thanks the ABIM Foundation and Eli Lilly and Company for their support for a workshop on clinician-level measurement the content of which contributed to the development of this Issue Brief.

NATIONAL QUALITY FORUM

601 Thirteenth Street, NW
Suite 500 North
Washington, D.C. 20005
202.783.1300 Tel
202.783.3434 Fax
www.qualityforum.org

Issue Brief No. 11
August 2008

© 2008. National Quality Forum

- 1 Gabel JR, Fahlman C, Kang R, et al., Where do I send thee? does physician-ownership affect referral patterns to ambulatory surgery centers? *Health Aff*, 2008;27(3):W165-W174.
- 2 See www.ncqa.org/tabid/59/Default.aspx.
- 3 See www.ncqa.org.
- 4 See www.ama-assn.org/ama/pub/category/2946.html.
- 5 Rattray MC, *Measuring Healthcare Resources Using Episodes of Care*; 2008. Available at www.aqaalliance.org/April3Meeting/materials/Measuring_Healthcare_Resources.pdf.
- 6 Rattray MC, Andrianos J, Stam DT, *Quality Implications of Efficiency-Based Clinician Profiling*; 2004. Available at www.carevariance.com/images/quality_Implications_Efficiency-Based_Clinician_Profiling.pdf.
- 7 Rattray MC, Andrianos J, Stam DT, *Revisiting Quality Implications of Efficiency-Based Clinician Profiling*; 2005. Available at www.carevariance.com/images/Revisiting_Quality_Implications_of_Efficiency-Based_Clinician_Profiling.pdf.
- 8 Safran DG, Karp M, Coltin K, et al., Measuring patients' experiences with individual primary care physicians, *J Gen Intern Med*, 2006;21:13-21.
- 9 See <http://healthcareDisclosure.org/docs/files/PatientCharter040108.pdf>.
- 10 Rattray MC, *Clinician Level Measurement and Improvement – Improving Reliability, Actionability, and Engagement*, Washington, DC: National Quality Forum; 2008.
- 11 See www.qualityforum.org/projects/ongoing/priorities/index.asp.
- 12 See www.qualityforum.org/pdf/projects/priorities/NQFFramework%20COMMENT%2006Nov07.pdf.
- 13 Ibid.
- 14 See www.brookings.edu/projects/qasc/collaborate.aspx.
- 15 Porter ME, Teisberg EO. *Redefining Health Care*. Boston: Harvard Business School Press; 2006.
- 16 See www.aqaalliance.org/files/PrinciplesofEfficiencyMeasurementApril2006.doc.
- 17 See www.ihl.org/IHI/Topics/Reliability/ReliabilityGeneral/ImprovementStories/GeisingerWarrantyonCABGSurgerySignalsCommitmenttoExcellence.htm.