Quality and Accountability: An Introduction for Advocates
Prepared by: David Machledt, PhD

Introduction: Why measure quality?

At first blush, assessing health care quality seems self-evident. Ideally, publicly available quality reports increase accountability, help reduce health disparities and aid consumers in health plan and provider selection.

However, the quality equation almost always includes a less heralded caveat – efficiency. This links care quality to cost. For example, to ensure that “efficiencies” do not derive from skimping on necessary care, provisions in the Affordable Care Act (ACA) tie performance to payment – such as withholding a percentage of payment until a provider or health plan meets certain quantifiable quality standards. Companies and organizations that develop and market quality measurement tools are now vying to establish their products as the standards.

But can quality measures alone provide an accurate picture of quality care? This Health Advocate outlines the promise and limitations of performance metrics. Consumers need measures that hold providers, health insurers and managed care entities (MCEs – risk-based and capitated plans and primary care case management systems) accountable to consistently deliver timely, patient-centered, coordinated care. But we are not there yet. Off-the-shelf measures may be too narrow, too easy to game or based on outdated or incomplete data. Moreover, the field of validated quality measures leaves significant gaps, such as Long Term Supports and Services (LTSS), care coordination and patient quality of life. Thus, an appropriately selected “family” of measures should always be accompanied by other assessment methods. In short, effective quality evaluation must be robust, multi-faceted and transparent.

Performance Measurement: Where are we now?

The predominant set of performance measures in health care today is the Health Effectiveness Data and Information Set (HEDIS®), published by the National Committee for Quality Assurance (NCQA), a private non-profit organization. Medicare, Medicaid, and many commercial ventures base their quality assessment programs on HEDIS measures.¹ Consequently, many providers already collect HEDIS data, which, in turn, reduces administrative burdens and makes HEDIS still more attractive. That said, not all health plans or providers with strong HEDIS results offer

¹ NCQA is also a key player in accrediting health plans for health marketplaces.
to top quality care, nor does HEDIS alone satisfy all Medicaid quality requirements (See Box 1). NCQA is not the only private organization participating in the performance measure market. Another example is the National Quality Forum (NQF), a private non-profit that evaluates and endorses performance measures. The number of NQF-endorsed measures has climbed sharply in the last ten years to nearly 700 in 2013.

Another source for quality measures is the family of Consumer Assessment of Healthcare Providers and Systems (CAHPS®) patient surveys developed by the Agency for Healthcare Research and Quality (AHRQ). CAHPS evaluates patient satisfaction with providers or health plans as well as other aspects of the care experience. AHRQ has compiled a large database with CAHPS performance benchmarks, but an Urban Institute review conducted for HHS found that, as with HEDIS, data collection methodologies varied considerably across states. This impedes effective comparisons across plans or states.

To improve comparability, numerous initiatives strive to select standardized measure sets for CHIP and Medicaid. CMS adopted and has since revised a [core measure set for children](https://www.cms.gov/Medicare/Quality-Improvement/QualityMeasures/Measures_CoreSets.html) as mandated by the CHIP Reauthorization Act of 2009. The ACA required CMS to develop [an initial core set for adults](https://www.cms.gov/Medicare/Quality-Improvement/QualityMeasures/Measures_CoreSets.html), published in January 2012. The goal of core measures is to increase uniformity and establish a mechanism to annually improve measurement techniques, but states’ participation is voluntary. In 2012, only Oregon collected data on all 24 child core measures, while 27 states reported on at least half. NQF has convened the “Measure Application Partnership” (MAP), a multi-stakeholder public-private partnership to select appropriate measures for evaluating public programs like Medicaid and Medicare. One aspect of the MAP initiative aims to endorse families of measures for individuals dually eligible for Medicare and Medicaid (“dual eligibles”). The [2012 MAP final dual eligibles report](https://www.nqf.org/ViewNQF/ProjectDetail.aspx?ProjectID=1796) also identifies gaps in currently available measures relevant to this diverse population with complex health needs (See Box 2.)

A Medicaid MCE’s performance can also be tracked through the state’s publicly available independent external quality reviews. Some external reviews go considerably beyond strict performance metrics and provide an in-depth analysis of plan performance in selected areas, using site visits, consumer focus groups and more.

### Characteristics of Good Measurement: Selecting an appropriate balance

An effective measure must demonstrably correlate with a desired quality outcome. Some measures provide useful and actionable data, but results often require interpretation and careful attention to context.

Broadly, performance measures consist of three types. **Structural** measures evaluate whether systems are in place, such as whether an MCE has developed guidelines for effective care. **Process** measures assess system function, like the percentage of patients receiving follow-up appointments after discharge from a hospital. **Outcome** measures monitor changes in beneficiary health status or behavior, such as the number of hospital readmissions within 30...
days of discharge. Generally, policy makers seek composite measures that would, for example, couple a metric for diabetes screenings (process) with evidence of follow-up treatment (process) and a measure of blood sugar control (outcome).

The following lists some key factors in the selection of an appropriate set of measures:

- **Specificity.** Some measures apply to subpopulations, such as patients discharged from an inpatient psychiatric setting, while others involve the entire population, such as screening for Body Mass Index. Overly specific measures can cause problems with sample size or lead MCEs to focus resources on a fraction of the population, while broad measures may either have too many variables to assess effectiveness or miss problems encountered by heavy care users.

- **Level of Analysis.** Measures can be designed to evaluate delivery systems, care settings, health plans, facilities, or individual clinicians.

- **Cost and Parsimony.** Providers and MCEs push states to limit the cost and the number of reported measures.

- **Inertia.** States tend to favor measures already in wide use to reduce administrative burden and, in the best cases, to increase comparability. To date, these have focused on structural and process measures.

- **Patient-centeredness.** Metrics should favor data reported by the actual users of the care and their families. Patient-centeredness should go beyond customer satisfaction, which can be colored by low expectations and other psychological factors, and focus on empowerment and autonomy, such as involvement in health care decision-making and access to self-directed care for people with disabilities.

- **Data consistency and reliability.** Some states collect data from administrative records (e.g. claims data), while others require both administrative data and reviews of medical records or allow MCEs to choose the method. These different reporting methods produce different, and incomparable, results.

- **Health disparities.** Measures should collect data stratified by key demographic categories (race, ethnicity, gender, socioeconomic status) to identify and target persistent health inequities.

- **Timeliness.** Some measures report results over a long term. While such results can be highly instructive, they may lag. A balanced measure set should also capture “real time” data, such as plan disenrollment statistics. Collecting data on grievances and complaints is a primary way to obtain real time information.

- **Stability.** Performance measures are most informative in showing long-term changes over a baseline. If MCEs or reporting requirements shift frequently, it becomes nearly impossible to generate actionable data. On the other hand, occasionally shifting or adding reported measures can discourage attempts to game the system and ensure that measurement reflects the state of the art.

---

**Box 2: Quality in Managed LTSS and Dual Eligible Demonstrations**

As part of an HHS initiative, 20 states have active proposals to redesign integrated care for dual eligibles. Participating states will implement payment withholds requiring MCEs to meet quality standards.

This initiative, coupled with new managed LTSS demonstrations in several states (KS, NJ, NY, FL), has highlighted key measurement gaps including care coordination, rebalancing from institutional to home and community-based care (HCBS), and quality-of-life. Advocates are pushing for more research to fill these gaps, but also stress that the dearth of appropriate LTSS quality measures heightens the importance of other assessment elements, like statistics on the LTSS direct work force (employee turnover, training and availability), MCE disenrollment data and measures of accessibility for people with disabilities. More detailed recommendations are available in NHeLP’s brief, “Five Key Standards for Dual Eligible MOUs.”

- **Cost and Parsimony.** Providers and MCEs push states to limit the cost and the number of reported measures.

- **Inertia.** States tend to favor measures already in wide use to reduce administrative burden and, in the best cases, to increase comparability. To date, these have focused on structural and process measures.

- **Patient-centeredness.** Metrics should favor data reported by the actual users of the care and their families. Patient-centeredness should go beyond customer satisfaction, which can be colored by low expectations and other psychological factors, and focus on empowerment and autonomy, such as involvement in health care decision-making and access to self-directed care for people with disabilities.

- **Data consistency and reliability.** Some states collect data from administrative records (e.g. claims data), while others require both administrative data and reviews of medical records or allow MCEs to choose the method. These different reporting methods produce different, and incomparable, results.

- **Health disparities.** Measures should collect data stratified by key demographic categories (race, ethnicity, gender, socioeconomic status) to identify and target persistent health inequities.

- **Timeliness.** Some measures report results over a long term. While such results can be highly instructive, they may lag. A balanced measure set should also capture “real time” data, such as plan disenrollment statistics. Collecting data on grievances and complaints is a primary way to obtain real time information.

- **Stability.** Performance measures are most informative in showing long-term changes over a baseline. If MCEs or reporting requirements shift frequently, it becomes nearly impossible to generate actionable data. On the other hand, occasionally shifting or adding reported measures can discourage attempts to game the system and ensure that measurement reflects the state of the art.
Conclusion: The need for transparency and a multi-faceted approach

Quality assessment is a permanent and growing feature on the health care landscape. The financial temptation for capitated MCEs to save money by reducing access to services only increases the importance of adequate oversight. While performance metrics should continue to improve over time, inconsistencies and gaps will always persist. A robust quality evaluation must combine appropriately selected quantitative metrics with other methods, like independent external reviews, reports on grievances and complaints and “secret shopper” surveys to evaluate real-time access to care.

Whatever the approach, quality measurement means little without transparency. Advocates cannot trust or evaluate results without knowing the context and methods of data collection. Because the performance measurement industry is mostly private, some data is considered proprietary or only available at considerable expense. Other data may be publicly available, but only through a lengthy public records request.

The Obama administration made an early commitment to increasing government transparency. The ACA required HHS to develop a National Strategy for Quality Improvement in Health Care through “a transparent collaborative process.” This national roadmap seeks broad stakeholder input to identify quality measures for federal health programs, set priorities for improvement, and promote coordinated data collection and reporting. An ongoing commitment to transparent collaboration is necessary to raise MCE and provider accountability and reinforce quality assessment as a critical consumer protection.