



Measuring Delivery System Performance — Toward a Common Denominator

Dhruv Khullar, B.A., and Dave A. Chokshi, M.D.

Mr. Khullar is from the Yale School of Medicine, and Dr. Chokshi from the Department of Veterans Affairs.

The recent formation of over 300 accountable care organizations (ACOs) reflects a belief in the potential for integrated health care delivery systems to promote the “triple aim”: better care for individuals, improved population health, and lower per capita costs. (<http://www.ncbi.nlm.nih.gov/pubmed/18474969>) In July 2013, the Center for Medicare and Medicaid Innovation (CMMI) released results from the first year of its Pioneer ACO pilot study, providing an important initial glimpse into the performance of these health systems. But truly capturing and comparing the overall value of these systems will require refining existing performance metrics and introducing new ones, as well as encouraging greater participation and standardization across a diverse set of payers and provider organizations. Improving our systems requires improving our measurements.

Various payers have deployed their own sets of evaluation criteria for hospitals and patients in ACOs, but there is currently no unified method to measure the performance of health systems as a whole. For example, while the Centers for Medicare and Medicaid Services and some other payers measure per capita costs, many do not, thus limiting the utility of these measures for making comparisons across systems with a mix of Medicare, Medicaid, and privately insured patients.

A number of factors have made risk-adjusted, comprehensive health system comparisons difficult. Barriers to data sharing have presented one significant obstacle, since privacy

and proprietary concerns, as well as technical challenges and market competition, have limited the extent to which health and cost data have been compared across systems. Of the comparison initiatives that do exist, most differ from holistic, system-level analysis in one or more important ways (see table).

First, many efforts focus on hospitals rather than health systems. Health systems that take responsibility for the spectrum of health needs of a population will necessarily span both inpatient and outpatient settings. In contrast, many current measurement efforts revolve around hospitals or hospital systems, such as those sponsored by the Leapfrog Group, Truven Health Analytics, HealthGrades, and the Joint Commission. An emphasis on hospital-based metrics fails to acknowledge the emerging shift of health care delivery from clinical facilities into communities.

Second, many studies use geography as the unit of analysis instead of health systems. The Commonwealth Fund has pioneered the development of rigorous scorecards to compare health system performance across local hospital referral regions, states, and countries. Applying a similar methodology to health systems themselves would enable more direct and actionable system improvement.

Third, lack of sophisticated risk adjustment has limited the possibility of apples-to-apples comparisons. While health plans have converged on risk adjustment methods such as Medicare’s Hierarchical Condition Categories calculation, there are no analogous harmo-

Pitfalls and Solutions for Comparing Health Systems.	
<i>Pitfall</i>	<i>Solution</i>
Focusing on hospitals instead of health systems: current efforts emphasize hospital-based performance metrics	Using a more comprehensive approach that incorporates inpatient and outpatient service performance measures
Using geography as the unit of analysis instead of health systems themselves	Instituting scorecards that compare integrated health systems directly
Lack of harmonized standards to compare health, cost, and quality outcomes across systems	Adapting risk-adjustment methodology from the health insurance industry to compare populations attributed to health systems

nized standards for health systems. As a result, there is no common denominator to compare health, quality, and cost outcomes. The need for more sophisticated risk adjustment driven by the creation of health insurance exchanges could help address this gap.

To account for the health status of the population enrolled in each insurance plan, the Affordable Care Act requires risk-adjustment mechanisms to reallocate premium income among plans. In March 2013, the Department of Health and Human Services finalized its risk adjustment methodology for the exchanges. With this methodology, payments can be adjusted according to the average risk score for a plan’s enrollee population, based on demographic characteristics and diagnostic information. An analogous methodology could be used to risk-stratify populations attributed to a given health system and allow for objective comparisons across health systems.

A first step toward system-level comparisons will be using measurements already recorded by health systems and selecting a standard set of metrics across which they can be compared, as CMMI did with the Pioneer ACO evaluation. The emphasis here should be on outcomes rather than process indicators—not simply whether a task is completed, but rather the extent to which a goal was realized. Several existing efforts can provide the foundation for quality of care delivery comparisons. For example, the Healthcare Effectiveness Data and Information Set and CMS

Physician Quality Reporting System are widely used databases measuring hospital and provider performance, and the Consumer Assessment of Healthcare Providers and Systems program uses patient surveys to evaluate patients’ experiences in healthcare settings. A carefully selected, consensus-based combination of these metrics could offer a path forward for care delivery comparisons across health systems.

A more complete evaluation of health system performance will require novel cost and population health metrics in addition to quality measures. One promising forum to this end is the Institute of Medicine (IOM) activity “Core Metrics for Better Care, Lower Costs, and Better Health.” (http://www.iom.edu/-/media/Files/Activity%20Files/Quality/VSRT/Core%20Metrics%20Workshop/Briefing%20Book_Combined.pdf) The consensus study seeks to harmonize methods for measuring health system performance based on the triple aim, including a set of population health metrics (e.g., length and quality of life, smoking rates, and utilization of preventive services) and cost measures (e.g., affordability, expenditures, and waste).

Building on these efforts, a complementary strategy for incorporating population health and spending goals might involve a single, accessible measure for each that can be easily compared across systems. An overarching population health measure such as health-adjusted life expectancy (HALE) is a promising metric that integrates morbidity and mortality

to express the number of years a person can expect to live in a healthy state. The U.S. Burden of Disease collaborators recently used HALE to summarize trends in the overall health of the U.S. population over time and compare these results to those of other OECD countries. (<http://www.healthmetricsandevaluation.org/gbd/publications/summaries/state-us-health-burden-disease-injuries-and-risk-factors>) A similar approach could be used to compare the overall health of patient populations across health systems within the United States. (<http://www.sciencedirect.com/science/article/pii/S2213076413000377>) A global measure such as HALE may also encourage recruitment and participation of broader health system stakeholders such as employers, housing agencies, and community groups to strengthen efforts to address underlying social determinants of health.

A similar summative measure of per capita costs might be used to evaluate the efficiency of health systems. It is important that such a metric include all relevant inpatient and outpatient expenditures, and that it be indexed for local market conditions. (<http://www.ncbi.nlm.nih.gov/pubmed/18474969>) A complete model would use the risk adjustment techniques described above to incorporate a host of popula-

tion and geographic characteristics including case-mix index, wage index, urban versus rural location, and poverty rate. (<http://www.ncbi.nlm.nih.gov/pubmed/19414903>) Implementation of electronic health records and health information exchanges will facilitate calculating and comparing risk-adjusted unit cost per patient across systems.

To improve health care delivery, its performance must be measured at the level where it can be changed. Initially, such measurement will depend on buy-in from a critical mass of health systems, ideally with the support of a convening body such as the IOM or National Quality Forum—that adopts the idea of common evaluation. A consensus-driven approach among leading health systems could forge a way ahead on thorny methodologic issues, such as standardizing risk adjustment, calculating real costs, and deciding upon time-frames for population health metrics.

But initial hurdles should not deter the march toward this important goal. Just as we look to comparative effectiveness research to make sound decisions about clinical interventions, we must seek comparative data on health system performance to continue to improve those systems.