Quality Measures for Dialysis: Time for a Balanced Scorecard

Alan S. Kliger

Abstract
Recent federal legislation establishes a merit-based incentive payment system for physicians, with a scorecard for each professional. The Centers for Medicare and Medicaid Services evaluate quality of care with clinical performance measures and have used these metrics for public reporting and payment to dialysis facilities. Similar metrics may be used for the future merit-based incentive payment system. In nephrology, most clinical performance measures measure processes and intermediate outcomes of care. These metrics were developed from population studies of best practice and do not identify opportunities for individualizing care on the basis of patient characteristics and individual goals of treatment. The In-Center Hemodialysis (ICH) Consumer Assessment of Healthcare Providers and Systems (CAHPS) survey examines patients' perception of care and has entered the arena to evaluate quality of care. A balanced scorecard of quality performance should include three elements: population-based best clinical practice, patient perceptions, and individually crafted patient goals of care.

Introduction
Measuring quality has become a central theme in United States health care. Clinical practice guidelines (CPGs), on the basis of best published clinical evidence and expert opinion, and clinical performance measures (CPMs) on the basis of those guidelines, are used as yardsticks to measure quality of care. Nephrology was among the first specialties to develop CPGs and CPMs, first in the United States with the Dialysis Outcomes Quality Initiative (1) and Kidney Disease Outcomes Quality Initiative (2), and later internationally with Kidney Disease Improving Global Outcomes (3). CPGs offer clinicians specific evidence-based care algorithms, designed to result in better patient outcomes. CPMs are clearly defined measures, specifying definitions of numerator and denominator to assess how well a population of patients adheres to a specific CPG (4). Selected CPMs have been used by the Centers for Medicare and Medicaid Services (CMS) and commercial insurers as measures of quality. When CPMs are reported publically, or are used for payment for services, they can exert a powerful influence on how we deliver care, with both intended and unintended consequences. In this article, I will review the ways in which CPMs have been used in nephrology, whether the metrics selected adequately reflect the quality of care, and consider the role of patient choice and patient-reported outcomes in measuring quality. I will then propose a pathway toward a more comprehensive and patient-centered process to measure quality for patients with kidney disease.

How Do We Use CPMs?
CPMs have been used for quality improvement efforts, public reporting, and payment.
who participate in the Physician Quality Reporting System program (12) and the Health Information Technology for Economic and Clinical Health Act promoting meaningful use (13). Payment for care is moving from fee-for-service to payment for performance (value-based purchasing) on the basis of CPMs. Physicians’ performance, similar to teachers’ performance, is increasingly measured not by the number of encounters or effort expended, but by the outcomes of those encounters. Just as teachers are judged by the test results of their students, doctors are increasingly judged, and paid, by the clinical outcomes of their patients. Clear evidence to support this trend can be found in recently passed legislation, the Medicare Access and Children’s Health Insurance Program Reauthorization Act of 2015 (14). This law, which repealed the sustainable growth rate as a basis for physician payment by Medicare, also establishes a merit-based incentive payment system (MIPS) and requires CMS to “develop a methodology for assessing the total performance of each MIPS eligible professional according to performance standards.” Performance will then be measured “using such methodology to provide for a composite performance score… (using a scoring scale of 0 to 100) for each such professional.” The MIPS will begin in 2019, and a 5% bonus will be available each year from 2019 to 2024 for physicians who agree to receive payment in alternative payment models, such as accountable care organizations. Beginning in 2026, fees will increase by 0.75% each year for providers who participate in an alternative payment model and 0.25% for those who do not. The law clearly incents physicians to move away from fee-for-service and to embrace pay-for-performance models of care. With the exception of the In-Center Hemodialysis (ICH) Consumer Assessment of Healthcare Providers and Systems (CAHPS) Survey, each of these metrics focuses on the processes of dialysis care or complications of the dialysis process. Are these the best

<table>
<thead>
<tr>
<th>Table 1. Centers for Medicare and Medicaid Services Dialysis Facility Compare quality measures: Five-star rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clinical Performance Measures</strong></td>
</tr>
<tr>
<td>1. Standardized hospitalization ratio</td>
</tr>
<tr>
<td>2. Standardized mortality ratio</td>
</tr>
<tr>
<td>3. Standardized transfusion ratio</td>
</tr>
<tr>
<td>4. Percentage of adult hemodialysis patients who had Kt/V urea ≥1.2</td>
</tr>
<tr>
<td>5. Percentage of pediatric hemodialysis patients who had Kt/V urea ≥1.2</td>
</tr>
<tr>
<td>6. Percentage of adult peritoneal dialysis patients who had weekly Kt/V urea ≥1.7</td>
</tr>
<tr>
<td>7. Percentage of adult dialysis patients who had an average calcium over the last 3 mo &gt;10.2 mg/dl (hypercalcemia)</td>
</tr>
<tr>
<td>8. Percentage of adult patients who received treatment through arteriovenous fistula</td>
</tr>
<tr>
<td>9. Percentage of adult patients who had a catheter left in a vein &gt;90 d for their regular hemodialysis treatment</td>
</tr>
</tbody>
</table>

The star rating combines the three dialysis adequacy (Kt/V) measures (nos. 4–6) into a single all Kt/V measure.

<table>
<thead>
<tr>
<th>Table 2. Centers for Medicare and Medicaid Services Quality Incentive Program for payment year 2018</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clinical Measures</strong></td>
</tr>
<tr>
<td>1. Vascular access: AV fistula: percent of patient months using an autogenous AV fistula with two needles</td>
</tr>
<tr>
<td>2. Vascular access: catheter ≥90 d: percent of patient months with a catheter used continuously for ≥90 d</td>
</tr>
<tr>
<td>3. Dialysis adequacy: adult and pediatric hemodialysis: spKt/V urea ≥1.2</td>
</tr>
<tr>
<td>4. Dialysis adequacy: adult peritoneal dialysis: weekly Kt/V urea (dialysis+residual) ≥1.7</td>
</tr>
<tr>
<td>5. Dialysis adequacy: pediatric peritoneal dialysis: weekly Kt/V urea (dialysis+residual) ≥1.8</td>
</tr>
<tr>
<td>6. Hypercalcemia: proportion of patient months with 3-month rolling average of total uncorrected serum Ca ≥10.2 mg/dl</td>
</tr>
<tr>
<td>7. NHSN bloodstream infection: hemodialysis: number of HD outpatients with positive cultures per 100 HD patient months</td>
</tr>
<tr>
<td>8. Standardized readmission ratio: unplanned readmissions: observed/expected</td>
</tr>
<tr>
<td>9. Standardized transfusion ratio: RBC transfusion events (one or more units): observed/expected</td>
</tr>
</tbody>
</table>

**Reporting Measures**

1. Mineral metabolism: number of months a facility reports serum phosphorus value for each Medicare patient
2. Anemia management: number of months a facility reports ESA dosage and Hb/Hct for each Medicare patient at least once per month
3. Pain assessment and follow-up
4. Clinical depression screen and follow-up
5. NHSN health care personnel influenza vaccination

AV, arteriovenous; sp, single pool; Ca, calcium; NHSN, National Healthcare Safety Network; HD, hemodialysis; RBC, red blood cell; ICH, In-Center Hemodialysis; CAHPS, Consumer Assessment of Healthcare Providers and Systems; ESA, erythropoiesis stimulating agent; Hb, hemoglobin; Hct, hematocrit.
measures of physician and dialysis facility performance? Do they focus at all on what patients and their families want from their renal replacement treatment? If CPMs are used to develop composite scores of physician performance, and those scores in turn are used to calculate physician payments, they become more than tools for quality improvement, they become driving forces determining the way care is provided. It is important therefore to evaluate carefully what aspects of care we are measuring, which CPMs we are using, and to design a quality evaluation plan for the future that measures the aspects of quality most important to our patients.

Which CPMs Are We Using?

The current set of kidney disease CPMs that aggregate patient data include outcome measures (mortality, hospital readmission, and patient experience of care), intermediate outcome measures (dialysis adequacy and vascular access), process measure (blood transfusion), and safety measures (blood stream infections and hypercalcemia) (Tables 1 and 2). Most (but not all) of these CPMs have been endorsed by the National Quality Forum (NQF). The NQF is a not-for-profit organization that fosters quality improvement and endorses consensus standards for performance measurement (15). By its own description, NQF endorsement is the gold standard for health care quality assuring that endorsed measures are evidence based and valid. Measure validity is reviewed at intervals, and measure developers are asked to document the adherence to endorsed standards over time. When adherence reaches very high levels, topped out standards are retired and replaced. Smith and Hayward note these measures do not individually provide a global assessment of quality (4). Intermediate outcomes must be clearly linked to clinically important outcomes to be meaningful measures. The integrity of such tight linkage over the range of reported values is needed for surrogate metrics to adequately represent the quality of care. In Nissenson’s review of quality metrics for patients with ESRD, the author argues that while quality improvement has been largely focused on biochemical/surrogate outcomes, there is an urgent need to move beyond these to focus on more patient-centered care (16). The author proposes a quality pyramid with its base-measuring biochemical and surrogate data, but moving up in a pyramid toward measures of effectiveness: outcome measures, including mortality, hospitalization, and patient experience measures. At the top of the pyramid the author proposes measures of health-related quality of life, what matters most to patients.

Hard outcome measures, such as mortality, may be most desirable, but defining useful and reproducible specifications for such measures can be challenging. For example, the standardized mortality ratio (SMR) is one of the publicly reported metrics for dialysis facilities. The SMR is a ratio of observed/expected deaths. There are several ways to calculate expected mortality rates (17). CMS also publicly reports hospital-level SMRs and risk-standardized 30-day mortality measures for acute myocardial infarction, heart failure, and pneumonia patients. If best care prolongs life, it seems self-evident that mortality is a key quality measure. However, several difficulties with this measure raise questions about its reliability. In one study, the several methods used to determine the SMR and results from these differences varied widely, calling into question the utility of the SMR (18). Another commentator points out that mortality as a hospital-level measure of quality presumes that death represents failure and discounts the fact that patients turn to hospitals for their end-of-life care (19). Holloway and Quill discuss mortality as a measure of quality and the implications for palliative and end-of-life care. They conclude that mortality is a good quality measure for individuals with acute illness who are not supposed to die, but a poor quality measure for most patients who suffer multiple chronic diseases and are near the end of their life (20). With an annual mortality rate of nearly 20%, dialysis patients often have multisystem disease, including cardiac, peripheral vascular, and metabolic diseases. For these patients, the overriding factors determining mortality may be largely these underlying disease processes, rather than the clinical management. Therefore, SMR is a blunt instrument, not patient centered because its interpretation is unclear for patients with limited life expectancy, and not specific to the dialysis care provided because comorbid factors are far more important than dialysis care.

One Set of CPMs for All Patients?

The observation that mortality may be a good quality indicator for healthy individuals but not for the chronically ill raises an important broader question for quality measurement: might some measures be appropriate for some patients and inappropriate for others? Patients have individual goals of therapy. Even when the disease process is similar, circumstances such as age, underlying comorbid conditions, frailty, predispositions, outlook toward life, life experiences, psychosocial anatomy and differences that we cannot adequately describe or measure determine how we each make choices in our health care. If we intend to deliver patient-centered care, should not these patient characteristics and choices determine what we seek for best quality care and quality measures be crafted around these patient-derived characteristics? Clearly, patients who come to dialysis as part of a palliative care plan have different goals than patients seeking life-prolonging treatment. Grubbs et al. define a palliative approach to dialysis care to prioritize comfort and align with patient preferences and goals of care most appropriate for patients in the last phase of their lives (21). For example, how do we best deliver and measure the quality of care for a 64-year-old woman with metastatic cancer who comes to dialysis after multiple chemotherapy treatments? She understands that her cancer is the determining feature of her health and her life expectancy is relatively short. She wishes to remain as symptom-free as possible, to have what time she has to help make plans for her elderly in-law, her cancer makes peritoneal dialysis impossible, and she requests in-center hemodialysis, with the proviso that her schedule be flexible enough to allow her to fulfill her goals of therapy. Is the measurement of Kt/V urea of any real relevance for her? She has small blood vessels and a primary arteriovenous (AV) fistula...
fails. Understanding the risk of possible infection, she requests no further surgery and wishes to remain on dialysis using a central venous catheter. Should we judge the quality of her care by the current CMS standard to maximize the use of AV fistulas and minimize the use of catheters? A patient-centered palliative approach to dialysis care might include the acceptability of central venous catheters, lower clearances, hypertension to avoid symptoms, and limited restrictions to diet (21). We may conclude that patients near the end of life seeking palliative care require a different set of quality metrics aimed more toward their short-term goals of therapy. Might there be other groups of patients who likewise require different quality metrics, for example frail older adults? How do we factor patient and family choices into our care plans and quality metrics?

The challenge is most pronounced when the quality metric is a process of care measure. For many years, the hemoglobin (Hb) or hematocrit was used as a surrogate measure of the quality of anemia management. Although Hb is one measure of anemia, it sometimes does not correlate with the symptoms or outcomes associated with anemia in individual patients. It has been argued (22) that patient-reported symptoms are a better measure of quality of anemia management than Hb level in some individuals. I cared for one man with stage 4 CKD and anemia who found that his energy level, ability to do usual daily activities, and ability to think at his best were impaired when his Hb fell <12 g/dL. On the basis of his own sense of wellness, he unerringly knew when his Hb fell below this threshold and requested erythropoiesis stimulating agent (ESA) therapy adjustment to maintain his health. He understood the clinical studies showing the risk of higher Hb levels for patients treated with ESAs, but he was clear that the risk was worth it to maintain his quality of life. Therefore, although population studies inform how quality metrics are best used for most members of that population, (e.g., best practice to maintain Hb<12), the best care plan for some individuals should be informed by factors such as patient choice and individual patient physiology. Wish points to this paradox between patient-centered care and quality metrics on the basis of population studies (23). Although AV fistulas are the preferred vascular access in the hemodialysis population, it may not be best for any given individual or in certain subsets of the population. Quality metrics must take both into consideration best treatment for populations and individual variation on the basis of patient-centered factors. Wish suggests an 80:20 rule to accommodate both (23). A goal should be set for a dialysis patient cohort to achieve 80% adherence to the population metric. For the 20% not achieving the goal, it is important to understand why. Individual patient choices and priorities, and their right to accept or refuse our recommended therapies, are legally guaranteed by the Patient Self-Determination Act of 1990 (24). Sometimes physicians will agree on goals of care for individuals deviating from the population standard. Other times, patients choose goals different than the physician advises. In either case, effort should be made to record the patient-specific goal of care and measure adherence to the goal.

Finkelstein suggests measuring patient-reported quality of life, while not lending itself to easily standardized assessments, should be the central focus of assessing care (25). The author argues each patient’s perception is unique and understanding how each patient experiences, understands, and adapts to the realities of ESRD care is needed to deliver patient-centered quality care. Therefore, measuring quality should likewise be patient centered, with some measure of what patients individually select as important to them. The challenge is resources are limited: if dialysis facilities and physicians spend time and money assuring dialysis adequacy, increasing AV fistulas, minimizing catheters, and reducing readmissions and blood transfusions to obtain their best five-star rating on Dialysis Facility Compare, little or no resources will be available for patient-centered measures.

**Goodhart’s Law**

Goodhart is an economist who examined the effect of regulation on a nation’s economy. He observed the following: “as soon as the government attempts to regulate any particular set of financial assets, these become unreliable as indicators of economic trends” (26). This concept has been generalized, asserting that when a measure becomes the target, it can no longer be used as a meaningful measure. For example, when standardized tests are used in the classroom to assess the quality of teaching, often the result is that teachers teach to the test to achieve adequate scores, but they do not otherwise change the quality of the teaching. Time and energy are spent tutoring students to answer the standardized questions correctly, while perhaps not teaching the logic, thought processes, and discipline needed to improve the quality of education. In part, the recent popular parent and teacher backlash against standardized testing in the classroom is the recognition that the metric no longer measures a child’s progress, but the time and attention required to prepare student for this examination is taking away valuable resources from more meaningful educational efforts. Goodhart’s law suggests that once the standardized test became the target of improving scores, it became an unreliable measure of educational progress. The same conundrum may exist for quality metrics in nephrology. Once a small set of measures is used to assess and pay for quality, those measures may cease to be meaningful quality metrics.

An additional challenge is to differentiate between a clinically important performance gap and a gap in documentation. In England, the National Health Service set quality goals to pay primary care doctors incentives for better quality care (27). They found the measures improved for diabetes and asthma care in the first years of the program. These improvements were on the basis of care reported in the medical records but not necessarily on care provided. The main effect of this pay-for-performance program may be to promote better recording of care rather than better care. We do not know to what extent CPMs in nephrology have improved as the result of better documentation versus true performance improvement. In current practice, physicians and dialysis facility owners work to bring the small number of CMS-designated measures to the required levels by focusing resources on those few measures, fixing record keeping, and achieving some real improvements in the measures. To the extent these
target measures have become unreliable measures of overall quality, similar to standardized tests in the classroom, this process may not change the real quality of care for dialysis patients.

Clinical Quality and Patient Experience Metrics

Although most nephrology CPMs measure processes and clinical outcomes, measures of the patient’s experience of care have entered the stage to evaluate the quality of care. CMS implemented the ICH CAHPS Survey in 2014 (28). Developed by the Agency for Healthcare Research and Quality, this tool measures the patient’s perception of care in the following six areas: (1) nephrologists’ communication and caring, (2) quality of dialysis center care and operations, (3) providing information to patients, (4) rating of kidney doctors, (5) rating of dialysis center staff, and (6) rating of dialysis center.

CMS plans to publicly report ICH CAHPS Survey results on the Dialysis Facility Compare website (9) on the basis of two semiannual administrations of patient surveys (29). The percentage of patient responses to this survey is also part of the Quality Incentive Program to determine payments to dialysis facilities. In England, where clinical quality metrics and patient experience have both been widely used to evaluate the quality of care, one study showed there were some associations between these two domains, but with very low correlations between clinical quality and interpersonal aspects of care (30). The authors suggest the quality of clinical care and quality of interpersonal care should be considered separately to assess overall medical care. Both are important: clinical quality metrics record how a patient’s care compares with best-evidence practice, and patient experience metrics record how delivered care compares with a patient’s best desired care. The ICH CAHPS Survey captures patient perceptions, particularly in the domain of interpersonal aspects of care. In the future, it seems likely CMS will use the scores for selected ICH CAHPS Survey domains to assess quality of care and will link these scores to payment. Once payers have experience using patient-reported quality of care metrics along with clinical quality metrics, these same measures might be the obvious choice to pay physicians in MIPS or an alternative payment model.

Summary

Our current system identifies several desired intermediate outcome and process measures and incent dialysis facilities to measure and improve these CPMs. Targeted efforts to improve these selected CPMs may limit attention to other quality needs, with the unintended consequence of diverting attention from individual patient quality gaps. Lack of adequate linkage between intermediate outcome and process measures with truly important clinical outcomes and Goodhart’s law may make targeted metrics unreliable measures of true quality. Biochemical/ surrogate outcomes may be relatively easy to abstract, aggregate, and report, but by themselves they are inadequate to reflect the overall quality of care. Measures of the patient experience and quality of life have recently been added to the renal CPMs, but to date they have only measured use of the ICH CAPHs tool, and not actual patient responses. CPMs and care standards have been one size fits all, with no differentiation of subpopulations with different quality needs, such as palliative care and frail elderly patients. Finally, our current CPMs do not capture patient choices, priorities, or patient and family values. Most of the current quality metrics are chosen by health care professionals on the basis of best clinical evidence, and not chosen by patients. Patients have different priorities and needs. Our current system has no way to craft individual patient quality goals—to define, measure, and improve specific patient-selected quality outcomes that matter to individual patients. Berwick says, “Every system is perfectly designed to achieve exactly the results it gets” (31). Our current system is perfectly designed to achieve certain population-based clinical process and outcome improvements. These were excellent goals for a start-up quality program. Now that our system is more mature, patient-centered, individual goal-directed care is more relevant, and our system should be changed to accommodate both goals.

Recommendations

1. Whatever their limitations, our current set of CPMs has functioned to improve specific aspects of ESRD patient care. Nissenson’s quality pyramid should be adopted by CMS, and more robust measures of effectiveness and quality of life should be developed and used for quality assessment (16).

2. CPMs to assess patient experience and engagement should be expanded. Moss and Davison suggest several such measures, including global symptom assessment, advance care planning, documentation of preferred surrogate decision maker in the event of incapacity, and referral to hospice when appropriate (32).

3. Patient-defined outcomes that matter to each individual should be crafted and used. We currently have no easy way to identify and measure such outcomes. Finkelstein (25) suggests that this requires exploration and study. The author suggests interdisciplinary care teams need to foster better internal communication and explore and discuss patients’ unique concerns. Better patient education, a curriculum for shared decision making for trainees, and making shared decision making part of routine clinical care will foster patient-centered care and lead to understanding of how such efforts can be measured. If dialysis facilities incented to support an infrastructure to determine patient-centered goals of care and selected appropriate, tested metrics for these goals, individual patient-selected CPMs could be developed.

4. MIPS and alternate methods for physician payment will be developed over the next decade. This offers us the opportunity to develop a balanced scorecard of CPMs for both physicians and dialysis facilities to include population-based clinical goals, patient experience measures, and individually crafted patient-reported goals. A more robust performance measurement system might include 50% CPMs, 20% patient experience measures, and 30% patient-defined and patient-reported measures. In this way, truly patient-centered care for patients with ESRD may be a model for accountable care and for
the alternative care models envisioned by the new federal legislation.

Disclosures
The author is a member of the National Quality Forum Renal Standing Committee.

References
2. Eknoyan G, Levin NW, Eschbach JW, Golper TA, Owen WF Jr, Schwab S, Steinberg EP; Baylor College of Medicine; Renal Research Institute; Beth Israel Hospital Center; Minor & James Medical Center; Vanderbilt University; Duke University Medical Center; Covance Health Economics and Outcomes Services; Johns Hopkins University; National Kidney Foundation’s Dialysis Outcomes Quality Initiative: Continuous quality improvement: DOQI becomes K/DOQI and is updated. *Am J Kidney Dis* 37: 179–194, 2001

Published online ahead of print. Publication date available at www.cjasn.org.