

# President's Page

## The Logic of Embracing Performance Measurement While Rejecting P4P

By Kenneth Y. Pauker, M.D., President



We are confronted with a quandary—how rationally to embrace the nascent science of performance measurement as a tool to improve quality in anesthesiology, while at the same time continuing to reject the flawed logic and self-serving rationale of Pay for Performance (P4P). This notion engenders in me—and I would venture in many of you as well—*cognitive dissonance*, “a state of psychological conflict or anxiety resulting from a contradiction between a person’s simultaneously held beliefs or attitudes.”

Dr. Steve Goldfien, Past President of the CSA, argues that this “performance assessment stuff” is unnecessary, unproven, and degrading to our profession. Dr. Bob Lagasse, former chair of ASA’s Committee on Performance Outcome Measurement (CPOM) and iconoclast at the ASA House of Delegates (HOD), bemoans the lack of scientific evidence and rigorous methodology behind the development of many performance measures to date. He further argues that the creation of what we now call performance measures inevitably leads to confusion about the nature of P4P, a similar-sounding but considerably more malignant scheme devised by allegedly well meaning—but mistaken and uninformed—cost- and quality-conscious health care policy wonks.

Dr. Alex Hannenberg, former ASA President, pushed hard to have the ASA accept P4P as more of an *opportunity* than a threat. But Dr. Mark Singleton, California’s ASA Director, the CSA Delegation, and I mounted a campaign against P4P at the 2007 ASA House of Delegates. Together, we were able to slow down the P4P train, at least temporarily, by suggesting two alternative approaches for anesthesiologists: *data collection* and *benchmarking*. Indeed, we brought forth a successful resolution to the ASA HOD, out of which evolved both the Anesthesia Quality Institute (AQI) and the National Anesthesia Clinical Outcomes Registry (NACOR).

The Perioperative Surgical Home concept (see pages 27–30) will benefit greatly from the massive clinical database that is being built. Measuring clinical outcomes adjusted for risk should, through the vehicles of benchmarking and creating practice parameters based on real science, advance the quality of anesthetic practice. This concept offers perhaps the last best hope for the

survival of anesthesiology amid the torrent of P4P measures that seem to be proliferating in an unrestrained manner.

### **Demands for More Care at Less Cost**

Business, government, insurers and patients all demand more care for less cost. The concept of P4P has led the Centers for Medicare and Medicaid Services (CMS) to develop “value based purchasing,” and we soon will be besieged with other methodologies of bundled payments such as accountable care organizations (ACOs) and “episodes of care.” Both hospitals and physicians are being coerced into participating in this evolving system of proving their own worth. Anesthesiologists are relative latecomers to the game, but payment restructuring looms ahead for us as well.

The real drivers of increased costs include:

- Demographics—both ends of the age spectrum
- The development of new and costly drugs
- The development of new technologies
- New, unproven diagnostic and therapeutic procedures
- Futile interventions near the end of life
- Defensive medicine

These are difficult to address because the nature of our social fabric abhors the idea of rationing care despite our finite resources, and because our political system is highly polarized and dysfunctional.

Certainly we are witness to many of the multitude of examples of unnecessary or unwarranted care. For instance, although not every back surgery is unwarranted, the evidence-based justification for many spinal fusions is being called into question. Moreover, the burgeoning of new but unproven procedures is hardly confined to spine surgery. Examples abound with joint replacements, cardiac and carotid stents, retinal therapies, robotic surgeries, CT scans and MRIs. Financial incentives for physicians and facilities encourage their use. As anesthesiologists, we often can be unwittingly complicit when we administer anesthetics that make unnecessary care possible. We live by the maxim that anesthesia is always necessary, even if the procedure may not be.

To be fair, we appreciate that there is a lack of evidence behind much of what we do as physicians, but also that this dearth of scientifically proven efficacy does not mean that we should not try new methods or procedures. What has become critical is that, in these times of careful cost scrutiny in medicine, the marginal benefit of deploying something new should certainly far outweigh its increase

in costs, such that true innovation is not stifled because it appears unnecessary or unwarranted to someone or some entity in authority.

On many levels, considerable energy is going into the eradication of unnecessary care as a way to control rising health care costs. Unfortunately, non-physicians may pick the wrong procedures or the wrong physicians to be the objects of their scrutiny. Physicians in general, and anesthesiologists in particular, are turning out to be easy targets.

### **Process, Structural, and Outcome Measures in Anesthesiology**

The practice of anesthesiology is different from other medical specialties. Despite careful planning and execution, we are often confronted with perturbations that require immediate judgment and decision-making with no time for rumination. How can performance measures take account of the competing priorities we must address in our most critical work? How can a perioperative beta-blocker be appropriate for a hemorrhaging and unstable patient? How can an outside reviewer balance the appropriateness of one course of action against another in a complex and rapidly changing situation? If performance measures are limited to processes only, they can't be done properly. However, *process* measures are just one of three broad categories of performance measures. There are also *structural* and *outcome* measures.

*Process* is what we do. It is concerned with appropriate and effective methods, done well. Guidelines and best practices, most familiar to anesthesiologists as substrates for measurement, are in this category. One recent example is the recent approval by the 2011 ASA HOD of Practice Guidelines for Central Venous Access, within which is a recommendation for the use of ultrasound. The 2010 HOD voted down a performance measure on using ultrasound, but with new guidelines, this will likely be revisited in 2012.

High-quality evidence is required for guidelines, and unfortunately, at present, an acceptable level of scientific evidence is often sorely lacking. "The use of recommendations [for performance measures] based solely on expert opinion or standards of care... lacks face validity, especially when such measures are to be used as the basis for public reporting or pay-for-performance."<sup>1</sup> Moreover, complex co-morbidities may call for competing or even mutually exclusive courses of action, and therein confound attempts to follow guidelines. In general, compliance with best practices in hospitals leads to better outcomes in only a modest way. CMS is now moving away from process measures and toward outcome measures in its Physician Quality Reporting System (see [www.cms.gov/PQRS](http://www.cms.gov/PQRS)).

*Structural measures* assume that “given the proper settings, good medical care will follow.”<sup>2</sup> Organizational characteristics, human resources, and technology are in this category. Examples include using state-of-the-art anesthesia machines or monitoring equipment, enhancing nurse or anesthesiologist staffing at designated times or for specific procedures, and implementing an electronic medical record or computerized order entry. Unfortunately, the evidence linking structure to outcomes is not robust. Nevertheless, insurers and regulators use this concept to incentivize information technology, despite the objections of both hospitals and physicians.

*Patient outcomes* are the final product of all clinical activity. In an individual patient, a specific outcome derives from a complex interplay of individual risk factors, chance, and effective medical care.<sup>3</sup> The quality of an outcome measure is a function of risk adjustment, data quality, sample size, and accurately choosing and then identifying appropriate outcomes. Examples are risk-adjusted mortality, perioperative myocardial infarctions and strokes, and long-term postoperative cognitive dysfunction.

With true risk-adjusted outcomes, eventually we will be able to benchmark one method against another. Then we will discover specifically what makes a significant difference to our patients. The fact that we have not yet arrived at that place with clinical data does not absolve us of our responsibility as physicians to try—with whatever tools we have—to improve how we care for patients. The problem for anesthesiologists is that others believe they already know now, and with unwarranted certainty, what quality is and how to measure it.

Anesthesiologists have been resting on our laurels for more than a decade. The Institute of Medicine commended us in 1999 as the one medical specialty that dramatically improved safety by a variety of approaches to reduce errors. However, Dr. Lagasse's analysis of data from multiple studies suggests that “anesthesia-related” mortality is much higher than the commonly quoted 1:250,000.<sup>4</sup> Clearly, we still have more work to do.

### **Distinguishing Performance Measurement from P4P**

In July 2007, the *ASA Newsletter* published a “Pro/Con” pair of articles on P4P.<sup>5</sup> Dr. Gerry Maccioli argued that an era of change, including incorporation of P4P models in compensation systems, was upon us, and that if we did not participate, then those who have no idea what we do would impose metrics upon us. The consequence of nonparticipation for our profession, he predicted, would be one of degradation and suffering. He proclaimed that P4P was more of an opportunity than a threat.

I argued, on the “Con” side, against the burden of P4P: its predictable new posse of tinkers; inevitable new ways to game a system with inadequate methods

for risk adjustment; rigid rule making; and a host of unintended consequences (such as paying large groups already in compliance, the avoidance of caring for high-risk patients, and the loss of access to care for disadvantaged patients). Four and one half years later, we witness what initially were small positive bonuses now giving way to increasingly significant negatives.

In my piece, I offered alternatives to P4P, some of which have come to pass:

*Payers (including the government) must invest in health information technology [payments for “meaningful use” are in process by the feds]. Specialty societies should mine the data (as with the ASA Closed Claims Project) and then set evidence-based standards and performance measures that enhance quality. Specialties could then construct benchmarks and provide detailed individual data to change individual behavior non-punitively [AQI and NACOR were constructed to do this]. Each specialty should demonstrate quality improvement, and there should be collaboration between specialties in areas of common interest and expertise [CPOM is beginning to take this approach]. Setting up a system of cutthroat competition between individual physicians within or across specialties will only exacerbate divisions within the House of Medicine...*

*Physicians and patients must be aligned and drawn together; intermediaries that pervert that relationship should be minimized and marginalized. Each specialty must address its unique inefficiencies and instances of profiteering and demonstrate this effort to payers. Further manipulations to squeeze additional “profits” and/or savings from already “pruned out” physicians must cease; instead, cost savings should be sought from “big ticket” items. These “big gorilla” items of waste and cost, such as outlandish health insurance industry profits, unnecessary and inappropriate care, and non-beneficial care at the end-of-life, must be debated and addressed by Medicine, society and government. We must tackle the thorny issue of what level of medical care our society can afford for all.*

Although we realize that performance measures ought to be constructed to improve quality, it is inevitable in this political and economic climate that others will use them to try to reduce what we are paid. I used to argue that for that exact reason we had to reject involvement in developing such measures at all. I've now become convinced that if we don't get involved, others—administrators, regulators and legislators—will construct them and impose them on us. They will use administrative instead of clinical data, and fail to employ scientific methods of risk adjustment. I see this in my own hospital where individuals charged with “continuous performance improvement” see only the surface of what anesthesiologists do. Nonetheless, they have specific directives from higher-level administrators to comply with hospital/corporate-centric interpretations of what CMS or The Joint Commission dictates.

### Measuring Performance Drives Quality Improvement

We must carefully translate what we know now into performance measures that will improve quality. We must do this now for our patients and ourselves, not merely to satisfy others. We must demonstrate that we know the way and will show the way. We are building NACOR and will benchmark ourselves. This will do more to improve quality and potentially reduce costs than anything else, but even this approach is unproven. Misadventures, misapplied technology, complications, unnecessary and nonbeneficial care, and defensive medicine will drive costs. Constructing better guidelines, understanding more efficient care, analyzing our individual modes of practice in relation to benchmarks—these will enhance safety, and as a byproduct, they will enhance value.

In some ways, Dr. Maccioli's 2007 analysis of what was coming was remarkably on target, but deploying performance measures to reduce payment for services remains a threat. On the other hand, measuring actual clinical performance, comparing it with established credible scientific criteria, and then monitoring for quality improvements in that performance, rather than focusing on cost and payment for meeting poorly or non-established performance parameters, is a worthy enterprise, and one befitting physicians whose priorities are science, safety, and enhancing quality in patient care.

What I have come to appreciate from my work on CPOM is this:

*What you don't know about how performance measures are constructed will hurt you.*

Understanding in depth how performance measurement is done, including its pitfalls and unintended consequences, is critical to being able to defend yourself from those who are more powerful and do not have your patients' or your best interests in mind. Arm yourself with knowledge to enable you to participate in the discussion, and prepare yourselves to champion science, logic, clarity and fairness in this new field of battle. Don't be a cog in the dissonance. Instead, embrace your cognitive dissonance, and stand ready with your colleagues in the CSA and ASA in advocating for performance measurement while continuing to reject P4P.

### References

1. Shaneyfelt TM, Centor RM. Reassessment of clinical practice guidelines: go gently into that good night. *JAMA* 2009; 301: 868–9
2. Donabedian A. Evaluating the quality of medical care: 1966. *Milbank Q* 2005; 83: 691–729
3. Glance LG, Neuman M, Martinez EA, Pauker KY, Dutton RP. Performance measurement at a “tipping point.” *Anes Anal* 2011; 112(4): 958–966
4. Lagasse RS. Anesthesia safety: Model or myth? *Anesthesiology* 2002; 97: 1609–17
5. Maccioli GA, Pauker KY. Pay for performance: pro and con. *ASA Newsletter* 2007; 71(7): 19–22