

Are We Overvaluing Performance Measures?

The title and some of the content of the editorial by Vijan¹ leave the impression that we are in danger of overvaluing performance measurement. However, there is, in fact, a dearth of scientifically sound, valid, and feasible clinical measures² available for use in medical practice. Thus, I have some concerns about the major points of the editorial.

Intermediate and Long-Term Outcomes Are Weakly Related

Other than a simple “cure” of acute pneumonia or similar event, most clinicians are aware of the many intervening variables that may impact “final” outcomes. However, a major finding of the diabetes study cited by Vijan was the substantial decrement in control over the 9 years of the UK Prospective Diabetes Study (UKPDS).³ Rather than indicating a problem with measurement, this would seem to relate to the need to continue both measurement and intervention. There is also the strong evidence from the Diabetes Control and Complications Study (DCCS)⁴ that good control of diabetes reduces the onset, and in some cases the progression, of a variety of micro- and macrovascular complications.

Dichotomous Measures Oversimplify the Relationship between Process and Outcome

Clearly, reducing a physiologic outcome measure such as hemoglobin (Hb) A_{1c} below some threshold, as well as the magnitude of change created, are both important to consider. In considering measures for clinical practice, there are always trade-offs in the way any measure is constructed or applied. In constructing Health Employer and Data Information Set measures for hypertension (blood pressure <140/90 mm Hg), low-density lipoprotein cholesterol after myocardial infarction (<130 mg/dL), and Hb A_{1c} (< 9.5%) in diabetes, we chose to use threshold for the following reasons. First, we set the threshold level such that *not* obtaining the level is strong evidence of inadequate control. Second, in both our measure development and subsequent field testing of the measures, given the large sample sizes used in HMOs (greater than 440), we found that commonly used risk adjustment had minimal effect on relative plan rank. Finally, using change, rather than threshold, introduced major problems in data collec-

tion and substantially increased the cost of data acquisition. These issues may not be critical where measures are used in an individual practice for internal quality-improvement purposes.

Universality Can Lead to Inefficiency

Finally, no one would dispute the issue that “universality can lead to inefficiency” or that “all patients are not the same.” However, it is also the case that the entire science base of medicine is built on our ability to classify individual patient illnesses into disease categories and to apply at least a core of standard treatments. Science-based and practical clinical performance measures are a critical part of the science base of effective clinical practice.

The issues I raise are directed at the inference that we “overvalue” performance measures. Despite the differences in some specifics, Dr. Vijan and I agree on the following: 1) Measurement alone is not sufficient; 2) we must measure what is important, not just what is easy to measure; and 3) not only an absolute level, but also the degree of positive change, should be examined.

The major difference between our viewpoints is that I believe we have paid too little, rather than too much, attention and devoted too few resources to creating clinical measures that address important diseases, that are based on careful scientific, and that are feasible to use in practice.

L. Gregory Pawlson, MD, MPH

Executive Vice President

National Committee for Quality Assurance

Washington, DC

References

1. Vijan S. Are we overvaluing performance measures? *Eff Clin Pract.* 2000;3:247-9.
2. McGlynn EA, Asch SM. Developing a clinical performance measure. *Am J Prev Med.* 1998;14:14-21.
3. UK Prospective Diabetes Study (UKPDS) Group. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). *Lancet.* 1998; 352:837-53.
4. The Diabetes Control and Complications Trial Research Group. The effectiveness of intensive treatment of diabetes on the development and progression of long-term complications in insulin dependent diabetes mellitus. *N Engl J Med.* 1993; 329:977-86.

THE AUTHOR RESPONDS

Although I have raised questions about performance measurement, they are not intended to discount its potential utility. Rather, these questions are meant to clarify the limitations of our current measures so that we can move forward to improve them.

The fact that intermediate and long-term outcomes may be weakly related must be considered when setting performance measures; the less strong the relationship, the less likely that measures in that area are worthwhile. Dr. Paulson cites the DCCT¹ and UKPDS² as showing that glycemic control reduces the onset and progression of microvascular disease. I completely agree with this point, although there is still no definitive evidence that improving glycemic control reduces macrovascular disease. Regardless, after 6 and 10 years of follow-up, only intermediate outcomes are affected in these two studies; changes in outcomes, such as visual acuity, renal failure, or mortality, have not been demonstrated. While improvements in intermediate outcomes will eventually lead to longer-term benefit, competing risk for mortality,³ especially in the older-onset patients with type 2 diabetes, will result in less improvement in outcomes that actually affect patients. This does not mean that we should ignore these areas; rather, the benefits should be quantified carefully and weighed against the costs of implementing quality improvement programs. Compare, for example, the benefits of hypertension control and glycemic control in diabetes. Studies of intensive hypertension control have shown dramatic and rapid benefit in symptomatic outcomes^{4, 5}; thus, performance measurement for aggressive hypertension control in diabetes should be a priority.

The decrement in glycemic control over the course of the UKPDS (despite a volunteer population and tremendous dedicated resources) is an important finding; however, rather than suggesting that more measurement would help, it raises the possibility that factors beyond the control of the health care system may prohibit us from achieving intensive glycemic control goals. As Dr. Paulson notes, good performance measures address this issue in part by setting looser standards (e.g., the target HbA_{1c} for quality monitoring is 9.5%, rather than the idealized 7.0%). However, this could be improved further. Measures that take into account changes in level of control, distance from goal, intensity of therapy, and case-mix or those that measure level of control across practices or systems (rather than providers)⁶ may be future methods (given current technical limitations) that will improve upon performance measurement and quality improvement. A closely related problem with our current performance measures is that they largely ignore patient preferences. This is troubling, as we may be imposing paternalistic standards requiring treatments that patients would

not want if they understood the associated benefits and burdens. Understanding variance in patient preferences, and how this variation should affect performance measurement, is another important area for future research.

As Dr. Paulson suggests, classifying patients is the first step toward minimizing the inefficiency of insisting that everyone receive the same level of treatment, regardless of risk. However, our methods of grouping can be improved, or even individualized, on the basis of risk.⁷ Few studies have specifically examined the proportion of benefit that is achieved by targeting high-risk populations; whether the incremental costs and benefits of broader-based interventions are worthwhile is uncertain in many cases. Understanding this is also an important part of the science of effective clinical practice.

Dr. Paulson seems to believe that I think that we pay too much attention to performance measurement. In fact, I feel exactly as he does. We need more study of how to implement quality measures that are proven to be effective and cost-effective and make sense in the context of both clinical practice and patients' lives.

Sandeep Vijan, MD, MS
University of Michigan
Veterans Affairs Medical Center Ann Arbor
Ann Arbor, Mich

References

1. The Diabetes Control and Complications Trial Research Group. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *N Engl J Med.* 1993; 329:977-86.
2. UK Prospective Diabetes Study (UKPDS) Group. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). *Lancet.* 1998;352:837-53.
3. Welch HG, Albertsen PC, Nease RF, Bubolz TA, Wasson JH. Estimating treatment benefits for the elderly: the effect of competing risks. *Ann Intern Med.* 1996;124:577-84.
4. UK Prospective Diabetes Study Group. Tight blood pressure control and risk of macrovascular and microvascular complications in type 2 diabetes: UKPDS 38. *BMJ.* 1998;317:703-13.
5. Hansson L, Zanchetti A, Carruthers SG, et al. Effects of intensive blood-pressure lowering and low-dose aspirin in patients with hypertension: principal results of the Hypertension Optimal Treatment (HOT) randomised trial. HOT Study Group [see comments]. *Lancet.* 1998;351:1755-62. Comment in: *Lancet.* 1998;351:1748-9.
6. Hofer TP, Hayward RA, Greenfield S, Wagner EH, Kaplan SH, Manning WG. The unreliability of individual physician "report cards" for assessing the costs and quality of care of a chronic disease. *JAMA.* 1999;281:2098-105.
7. Rothwell P. Can overall results of randomized trials be applied to all patients? *Lancet.* 1995;345:1616-9.