

10. Shrestha MP, Scott RM, Joshi DM, et al. Safety and efficacy of a recombinant hepatitis E vaccine. *N Engl J Med* 2007;356:895-903.

11. Jothikumar N, Aparna K, Kamatchiammal S, Paulmurugan

R, Saravanadevi S, Khanna P. Detection of hepatitis E virus in raw and treated wastewater with the polymerase chain reaction. *Appl Environ Microbiol* 1993;59:2558-62.

Copyright © 2007 Massachusetts Medical Society.

Performance Measurement in Search of a Path

Rodney A. Hayward, M.D.

In this issue of the *Journal*, Landon and colleagues report on a Herculean undertaking — a study of quality-improvement interventions conducted at 44 community health centers.¹ This study showed a modest improvement in some process measures and no improvement in intermediate or end-stage outcomes — results that are similar to those of most previous large-scale quality-improvement initiatives. As the authors correctly note, improved processes may not be accompanied by discernible improvements in outcomes for several reasons. A particularly important problem is that the majority of Health Plan Employer Data and Information Set-style performance measures used to guide most large-scale quality-improvement activities represent inefficient and sometimes counterproductive standards for improving clinical outcomes.

The shortcomings of current approaches to performance measurement are distressing, since it may be the single most important health policy tool for improving health care. Our experience with performance measurement over the past two decades has generally shown that “what you measure improves,” but unfortunately, we often settle for measuring that which is simple and easy to gauge and then sit back and celebrate the improvements in our “measures.” As a result, we risk wasting both resources and opportunities. With the current interest in pay-for-performance programs, this is an opportune time to revisit the principles and rationale underlying meaningful performance measurement.

First, it is critically important to understand that performance measurements are inherently and fundamentally different from clinical guidelines.²⁻⁴ In a guideline (an educational tool designed to aid clinicians in providing optimal care), it might be perfectly acceptable to recommend glycated hemoglobin levels of less than 7% (or even <6.5%) as a goal or to recommend annual eye screening for patients with diabetes. However, for many reasons that have been chronicled over

the years,²⁻⁷ basic guidelines are rarely appropriate as “all-or-nothing” performance measures. The reasons that guidelines often make poor performance measures are nonintuitive and easily forgotten by those who do not take care of patients. Indeed, in the very political and high-stakes process of selecting performance measures, influential parties often have strong incentives to advocate that these measures be aligned with idealized goals. For example, even the most pure-hearted persons and groups with a vested interest in issues related to diabetes (such as the American Diabetes Association, diabetes “experts,” and the diabetes sections of the National Institutes of Health and the Centers for Disease Control and Prevention) have a natural and justifiable tendency to want more attention and resources for their cause and will logically want to push for the most care for patients with this disease.⁸ It sounds terrible when we hear that 50% of recommended care is not received, but much of the care recommended by subspecialty groups is of modest or unproven value, and mandating adherence to these recommendations is not necessarily in the best interest of patients or society. For example, it has been estimated that it would take a primary care physician almost 8 hours per workday just to provide the interventions recommended by the U.S. Preventive Services Task Force, leaving no time to deal with acute conditions.⁹

At the heart of this problem is our wish to keep efforts at quality improvement and cost containment separate. When selecting quality measures, leaders of the performance-measurement process nationally tend to ignore the issues of the burden imposed on patients, patients’ preferences, and costs, preferring instead to construct a separate cost “report card.” One leader in a national performance-measurement organization told me that it is irrelevant how much benefit accrues from performance measurement, “just as long as it’s recommended by a respected professional group

and the benefit is not zero.” The current paradigm simply replicates the conventional pressures within the health care system, with the performance measures for quality suggesting that clinicians should do everything for everyone, regardless of costs or patients’ preferences, and the measures for cost-efficiency indicating that it is imperative to keep costs down — producing strong incentives not to care for sick people — and to skimp on or decrease coverage for care that is not part of the report card.

Rather than oscillate between berating providers for soaring health care costs and decrying any deviation from speculative recommendations, we need to come to grips with the fundamental problem in U.S. health care — we get exceedingly poor value from the dollars that we spend. Toward this end, performance measurement should be used as a tool to help us better direct our attention and resources where they will do the most good. As long as incremental dollars continue to go to low-priority care as often as they go to high-priority care, the tragic condition of U.S. health care will continue to worsen. If you think that the current situation is not a tragedy, consider that health care accounts for more than \$1 of every \$7 spent in the U.S. economy, which is much more than in any other country, yet more than 40 million U.S. citizens lack basic health care coverage, and our most vulnerable citizens — the sick, the indigent, and the elderly — are being saddled with devastating health-related financial burdens because of inadequate insurance. The last thing we need is a performance-measurement system that encourages a little improvement in quality and substantial increases in costs.

Performance measurement can encourage greater value in health care with one of two approaches. The easier way is to construct simple measures of high-priority care such as screening for diabetic retinopathy at least every other year¹⁰ and prescribing at least moderate-dose statins for patients at high risk for coronary artery disease.¹¹ The other approach is to encourage optimal care by using weighted measures that give more credit for high-priority care than for lower-priority care. However, if we wish to try to measure optimal care, we will almost always need to construct performance measures that are much more nuanced and that consider patients’ preferences, competing needs, and the complex circumstances of individual patients.^{2-6,12-15} Extensive work has shown

how simplistic, all-or-nothing performance measurement can mislead providers into prioritizing low-value care and can create undue incentives for getting rid of “bad” patients.^{3,5-7,10,11,16}

Finally, perhaps our greatest barrier to developing a worthwhile performance-measurement system is our unwillingness to invest in it. If we wish to look at what can be accomplished by investing in a clinically meaningful performance-measurement system, we can learn much by examining the experience of the Department of Veterans Affairs over the past 10 years. There is reason to believe that the agency’s celebrated rise to international leadership in quality is largely related to its investment in performance measurement.¹⁵ Despite having one of the most sophisticated integrated, electronic medical-record systems in the world, the Department of Veterans Affairs still invests heavily in a detailed review of medical records when constructing performance measures. It is pure denial to believe that we can untangle the good from the bad and the wasteful in medicine without clinically detailed information. The value and importance of most medical treatments vary tremendously among patient populations in complex ways. Therefore, the clinical context is essential to distinguish care of high value from more discretionary and contraindicated care.^{2-7,10-14} Until our performance-measurement system is based on clinically relevant information and targets high-priority care, performance measurement is likely to remain a great idea that is more of a distraction than a benefit.

No potential conflict of interest relevant to this article was reported.

The views expressed in this editorial are those of the author and do not necessarily represent the views of the Department of Veterans Affairs or the University of Michigan.

From the VA Ann Arbor Health Services Research and Development Center of Excellence and the Schools of Medicine and Public Health, University of Michigan — both in Ann Arbor.

1. Landon BE, Hicks LS, O’Malley AJ, et al. Improving the management of chronic disease at community health centers. *N Engl J Med* 2007;356:921-34.
2. O’Malley AS, Clancy C, Thompson J, Korabathina R, Meyer GS. Clinical practice guidelines and performance indicators as related — but often misunderstood — tools. *Jt Comm J Qual Saf* 2004;30:163-71.
3. Hayward RA, Hofer TP, Kerr EA, Krein SL. Quality improvement initiatives: issues in moving from diabetes guidelines to policy. *Diabetes Care* 2004;27:Suppl 2:B54-B60.
4. Brook RH, McGlynn EA, Shekelle PG. Defining and measuring quality of care: a perspective from US researchers. *Int J Qual Health Care* 2000;12:281-95.
5. Kerr EA, Smith DM, Hogan MM, et al. Building a better quality measure: are some patients with ‘poor quality’ actually getting good care? *Med Care* 2003;41:1173-82.

6. Kerr EA, Krein SL, Vijan S, Hofer TP, Hayward RA. Avoiding pitfalls in chronic disease quality measurement: a case for the next generation of technical quality measures. *Am J Manag Care* 2001;7:1033-43.
7. Hofer TP, Hayward RA, Greenfield S, Wagner EH, Kaplan SH, Manning WG. The un-reliability of individual physician "report cards" for assessing the costs and quality of care of a chronic disease. *JAMA* 1999;281:2098-105.
8. Deyo RA, Patrick DL. Hope or hype: the obsession with medical advances and the high cost of false promises. New York: AMACOM, 2005.
9. Yarnall KS, Pollak KI, Ostbye T, Krause KM, Michener JL. Primary care: is there enough time for prevention? *Am J Public Health* 2003;93:635-41.
10. Hayward RA, Cowan C Jr, Giri V, Lawrence MG, Makki F. Causes of preventable visual loss in type 2 diabetes mellitus: an evaluation of suboptimally timed retinal photocoagulation. *J Gen Intern Med* 2005;20:467-9.
11. Hayward RA, Hofer TP, Vijan S. Narrative review: lack of evidence for recommended low-density lipoprotein treatment targets: a solvable problem. *Ann Intern Med* 2006;145:520-30.
12. Pogach LM, Rajan M, Aron DC. Comparison of weighted performance measurement and dichotomous thresholds for glycemic control in the Veterans Health Administration. *Diabetes Care* 2006;29:241-6.
13. Hayward RA, Kent DM, Vijan S, Hofer TP. Reporting clinical trial results to inform clinical providers, payers and consumers. *Health Aff (Millwood)* 2005;24:1571-81.
14. McMahon LF Jr, Hayward RA, Saint S, Chernew ME, Fendrick AM. Univariate solutions in a multivariate world: can we afford to practice as in the "good old days"? *Am J Manag Care* 2005;11:473-6.
15. Asch SM, McGlynn EA, Hogan MM, et al. Comparison of quality of care for patients in the Veterans Health Administration and patients in a national sample. *Ann Intern Med* 2004;141:938-45.
16. Hofer TP, Zemencuk JK, Hayward RA. When there is too much to do: how practicing physicians prioritize among recommended interventions. *J Gen Intern Med* 2004;19:646-53. [Erratum, *J Gen Intern Med* 2004;19:903.]

Copyright © 2007 Massachusetts Medical Society.