

Is It Defensible to Use Volume Standards for Purchasing Care?

Lucian L. Leape, MD

The emergence of the Leapfrog group initiative to use measures of quality and safety as criteria for contracting for provision of medical services was a watershed event. Major payers would pay for quality, or, more accurately, no longer pay for inferior quality. Not surprisingly, physicians have been skeptical about this enterprise. Battered over the past 2 decades by an never-ending array of purchaser/payer/plan restrictions on medical practice that were clearly designed to save money (and increase profits), many saw this as yet 1 more device—a ruse, perhaps—for cutting costs—at doctors' expense. Equally galling was the use of volume of services as a measure of quality. For them, the findings by Goodney et al that there are no significant differences overall in length of stay (LOS) or return to the hospital between high-volume and low-volume hospitals comes as welcome news: See, it doesn't save money after all.

But ironically the results are also pleasing for those who welcomed this innovative approach to improving quality of care: See, higher quality doesn't cost more. In fact, there are several aspects of the study that suggest these measured differences don't tell the entire cost story.

First, deaths in the operating rooms obviously reduce lengths of stay. To the extent that mortality differences result from more of these early deaths, LOS would be lower in low-volume hospitals. If so, increased costs in high-volume hospitals would be worth paying for. Second, academic health centers tend to be high-volume hospitals (one of the reasons there has been little outcry about the Leapfrog requirements from academia). Both costs and lengths of stay are often higher in teaching hospitals, and for complicated patients, quality of care is better too. Unless data are adjusted for this variable, the results could be skewed. Finally, the use of logarithmic transformation to deal with the problem of rare very long stays, though justified for performing regression analysis, could also skew the data. Very long lengths of stay can have a profound impact on the average LOS. If low-volume hospitals with more complications also have a higher fraction overall that have very long stays (a reasonable and testable hypothesis) reducing the impact of those cases by logarithmic transformation favors low-volume hospitals.

But, if we accept this current news that there are no significant difference overall in length of stay and returns to the hospital at face value, as well as the authors' important prior work that demonstrated better outcomes in high-volume hospitals, it is still legitimate to question whether volume of cases is a suitable measure by which to judge quality. Most of us would say it is not. Surgeons, far more than the public, the policy makers, and even our medical colleagues, recognize the profound effect of surgical skill on outcomes, as well as the remarkable variation of its distribution among our colleagues. We also know that these differences are only modestly related to the volume of procedures

From the Harvard School of Public Health, Boston, MA.

Reprints: Harvard School of Public Health, 677 Huntington Ave., Boston, MA 02115. E-mail: leape@hsph.harvard.edu.

Copyright © 2003 by Lippincott Williams & Wilkins

0003-4932/03/23802-0168

DOI: 10.1097/01.sla.0000081084.36358.2d

performed by the individual. Although the answer to the question, does practice make perfect or does perfect make practice, may be “both”, we all know some surgeons who get good results with a modest number of cases for less common operations.

Volume of cases is but a surrogate measure of quality, and, as such, a very blunt instrument. The better alternative is to measure actual individual or hospital outcomes, with appropriate case-mix and risk-adjustment. And to measure all outcomes, not just death and complications, but relief of symptoms and return to full function as well. Fortunately, methods for doing this are improving every year. The experience with the Veterans Health Administration National Surgical Quality Improvement Project demonstrates the usefulness of this approach.¹ As surgeons, it is our obligation—and, now, in the new world of profiling, report cards, and disclosure, very much in our interest—to lead and participate in studies to refine and validate risk-adjustment techniques.

Once we have valid and reproducible methods for risk adjustment, we should press for their use instead of crude

surrogates such as volume. There is reason to believe that the Leapfrog group would welcome this alternative. Most of us would prefer to go, as a patient, to a surgeon with a risk-adjusted 1% mortality for a given operation over one with a 3% mortality, regardless of volume (or LOS or rate of return to hospital, for that matter).

Many of the arguments against use of volume guides for purchasing care, however, have not focused on quality, but on the secondary effects of such a policy, such as the threat it represents to the already precarious supply of some specialty surgeons in rural areas and to small hospitals, generally. Shifting to a risk-adjusted outcome approach would not solve those problems. Creative thinking is needed to develop better methods for maintaining access to specialty care in such situations. Settling for lower quality, however measured, should not be one of them.

REFERENCE

1. Khuri SF, Daley J, Henderson WG. The comparative assessment and improvement of quality of surgical care in the Department of Veterans Affairs. *Arch Surg*. 2002;137:20–27.