

Does Practice Really Make Perfect?

Surgery is risky. Worries include surgical errors, perioperative complications, difficult recovery, and poor long-term outcomes. Given the high stakes of surgical treatment, where should patients go when they need surgery? Common sense and a growing literature suggest that they should seek surgeons and hospitals that frequently perform the needed procedure. After all, practice makes perfect, doesn't it? Two studies reported in this issue suggest that the answer to this question is considerably more complicated than it appears (1, 2).

In raising questions about volume–outcome relationships, these studies swim against a strong current. Corporate America embraces the notion that practice makes perfect. Organizations that evaluate health care offer profiles that feature volume prominently (3–5). The Leapfrog Group, a coalition of more than 140 companies convened to improve the quality of health care for their employees, believes that the evidence links high procedure volumes to favorable outcomes and recommends specific volume thresholds for a variety of procedures (6). Many who study health care delivery also believe that high volume means superior outcomes. Researchers analyzing the potential benefits of the Leapfrog Group's volume standards for 5 high-risk surgical procedures estimate that the requirements would save 2581 lives annually (7). A reviewer of 20 years of volume–outcome studies concluded that high volume is associated with better outcomes across a variety of conditions (8).

Several potential explanations lie behind past conclusions that surgical patients fare best with high-volume providers. The first is that practice really does make perfect. To test this hypothesis, we ask whether patients who underwent surgery at low-volume hospitals or by low-volume surgeons would have done better had they gone to high-volume providers. A definitive answer to this question would require a randomized trial, but only observational data exist. Observational studies of volume–outcome relationships examine aggregated data derived from groups of patients treated at hospitals that are categorized by surgery volume status. Policymakers and others have concluded from these studies that redirecting individual patients from low-volume providers to high-volume providers will improve their outcomes. However, such inferences are an ecological fallacy, a false inference about individual outcomes based on aggregate data obtained from groups. To make correct inferences, we must carefully consider multiple explanations for volume–outcome associations: patient factors, provider factors, and analytical techniques.

PATIENT FACTORS

High-volume providers might achieve favorable outcomes because patients who seek their care are generally healthier or otherwise more advantaged than patients who

seek care from low-volume providers. Arguments against this explanation note that volume–outcome relationships persist after adjustments for comorbidity. However, simply adjusting for the presence of comorbidity may be insufficient, as suggested by the following example. Two patients have identical-stage pancreatic cancer, type 2 diabetes, and hypertension. Patient A presents to hospital A, a hospital that averages 3 Whipple procedures a year. Patient B goes to hospital B, a hospital that averages 50 Whipple procedures per year. Hospital A is in a poor, urban neighborhood, and many of its patients (like patient A) are poorly educated, unemployed problem drinkers with stingy insurance. Patient A takes his antihypertensive agents and diabetes agents sporadically when he can afford them and is not on a drinking binge. Patient B (like most of hospital B's patients) is a well-educated, employed teetotaler with generous insurance. Patient B takes his medications religiously. While a study comparing outcomes among patients treated at hospitals A and B might adjust for the presence of comorbidity (hypertension and diabetes) and insurance status, most volume–outcome studies do not adequately account for the more subtle sociodemographic differences among the hospitals' patient populations that may confound associations between volume and outcomes. Furthermore, adjusting for the presence of comorbidity is not the same as adjusting for the quality of care for that comorbidity. For example, diabetic patients who undergo surgery at high-volume hospitals might receive excellent outpatient diabetes care and have glycemic control, while those at low-volume hospitals have poor care and control. Rather than expertise, fundamental differences between the patients of low- and high-volume providers may account for the more favorable outcomes observed among high-volume providers. Indirect evidence for this explanation lies in studies of more homogeneous populations, such as a study of all patients who underwent 1 of 5 types of surgery at Veterans Affairs medical centers (9), which tend to find more equivocal volume–outcome relationships than do studies of heterogeneous populations.

In the first of the 2 volume–outcome studies in this issue, Meyerhardt and colleagues (1) demonstrate that patient factors can make care at low-volume hospitals appear worse than care at high-volume hospitals. In a study of 3161 colon cancer resections, these investigators found that patients who had cancer resected at low-volume hospitals had worse long-term survival than patients at high-volume hospitals. However, patients at low-volume hospitals did not have higher rates of death from colon cancer. In fact, cancer recurrence and cancer deaths did not vary with volume. Patients at low-volume hospitals died sooner than those at high-volume hospitals not because poor surgical technique led to cancer recurrence and subsequent death but rather because of conditions unrelated to colon cancer

or its surgical treatment. This study offers 2 important lessons about volume–outcome studies. First, one must examine condition-specific outcomes that the quality of the surgery is likely to influence in order to show that differences in outcomes are due to differences in surgical quality. Second, patient factors rather than surgical expertise may account for the observed differences between high- and low-volume providers.

PROVIDER FACTORS

Provider factors that happen to be more common among high-volume providers may mislead observers to conclude that volume per se drives outcomes. Sophisticated hospital computer systems, high nurse–patient ratios, or state-of-the-art anesthesia equipment might improve surgical outcomes in any hospital yet could be more common in high-volume hospitals. For example, a study of 10 complex procedures in U.S. hospitals found that higher staffing levels for resident physicians and nurses at high- compared with low-volume hospitals accounted for observed links between high volumes and favorable outcomes (10). Low-volume hospitals with staffing levels similar to those at high-volume hospitals achieved outcomes on par with those of high-volume hospitals.

ANALYTICAL TECHNIQUES

The presence and magnitude of volume–outcome associations may depend on the analytical techniques investigators use to examine these associations. Volume–outcome studies typically test the hypothesis that outcome rates differ by more than expected by chance and often rely on statistical tests that assume the outcomes of individual patients are independent of each another. However, for many reasons, a hospital or surgeon tends to attract similar patients. This natural “clustering” means that 2 patients of a single hospital (or surgeon) are more likely to be similar than are 2 patients of different hospitals (or surgeons) (11). In the presence of clustering, researchers must study more patients to be reasonably certain that there are true differences related to volume. Researchers must also use statistical techniques that take clustering into account rather than methods that assume independence between individual patients of a given provider. If they do not, they risk concluding that volume is the cause of differences between providers when the observed differences are actually consistent with random variation.

In the second volume–outcome study reported in this issue, Panageas and colleagues (2) show that inattention to clustering could lead to potentially erroneous conclusions about the link between volume and outcome. They measured associations of provider’s volume of surgical procedures and surgical outcomes for 3 types of cancer surgeries with and without adjusting for clustering of outcomes by surgeon. In some cases, adjustment reduced the statistical significance of volume-related outcome differences (as re-

flected by widening confidence intervals for the odds of complications). A reduction in statistical significance with adjustment for clustering indicates that the outcomes of individual patients of a provider were not independent. Panageas and colleagues also demonstrate that different statistical methods for adjusting for clustering give rise to different estimates of the size of the volume–outcome effect. In fact, the best way to perform statistical analysis of volume–outcome relationships is far from settled (12). Only further methodologic research will determine the ideal approach to these issues.

From Panageas and colleagues we learn to be skeptical of claims of “statistically significant” effects of volume on outcomes unless the authors have adjusted for clustering. We also learn that clustering of outcomes by provider can offer clues to variations in quality of care. Probing the practices of provider clusters with good outcomes (regardless of their procedure volumes) can provide insight into factors that promote desired outcomes.

CONCLUSION

While practice surely facilitates progress along the path to better surgical outcomes, its importance relative to other factors is far from clear. Differences in patient and provider factors can make volume appear to have more influence on outcomes than it really does. Because many studies have neglected clustering, failed to focus on condition-specific outcomes, inadequately adjusted for confounders, or assumed that aggregate outcome data predict the outcomes of individuals, the relationships between volume and outcome are neither as straightforward nor as robust as much of the literature suggests. The 2 studies in this issue examine volume–outcome relationships for surgical conditions, but the same arguments apply to volume–outcome studies of medical conditions such as HIV infection (13, 14). In fact, studies of medical conditions may be more complicated than studies of surgery because episodes of care are difficult to define, multiple providers often care for a single patient, and outcomes develop over a protracted period. “Does practice make perfect?” is a question that we cannot answer easily. Unfortunately, it appears that perfecting outcomes will require more than simply directing patients to providers with the most cases.

Christine Laine, MD, MPH
Senior Deputy Editor

Harold C. Sox, MD
Editor

Acknowledgments: The authors thank Drs. Paul Epstein, Eliseo Gualtar, Russell Localio, and Cynthia Mulrow for their careful review and helpful comments during preparation of this editorial.

Requests for Single Reprints: Customer Service, American College of Physicians, 190 N. Independence Mall West, Philadelphia, PA 19106.

Ann Intern Med. 2003;139:696-698.

References

1. Meyerhardt JA, Catalano PJ, Schrag D, Ayanian JZ, Haller DG, Mayer RJ, et al. Association of hospital procedure volume and outcomes in patients with colon cancer at high risk for recurrence. *Ann Intern Med.* 2003;139:649-57.
2. Panageas KS, Schrag D, Riedel E, Bach PB, Begg CB. The effect of clustering of outcomes on the association of procedure volume and surgical outcomes. *Ann Intern Med.* 2003;139:658-65.
3. DoctorQuality, Inc. Accessed at www.doctorquality.com on 29 July 2003.
4. HealthScope. Accessed at www.healthscope.org on 29 July 2003.
5. Select Quality Care. Accessed at www.selectqualitycare.com on 29 July 2003.
6. The Leapfrog Group. Accessed at www.leapfroggroup.org on 17 July 2003.
7. Birkmeyer JD, Finlayson EV, Birkmeyer CM. Volume standards for high-risk surgical procedures: potential benefits of the Leapfrog initiative. *Surgery.* 2001;130:415-22. [PMID: 11562662]
8. Halm EA, Lee C, Chassin MR. Is volume related to outcome in health care? A systematic review and methodologic critique of the literature. *Ann Intern Med.* 2002;137:511-20. [PMID: 12230353]
9. Bates EW, Berki SE, Homan RK, Lindenaue SM. The challenge of benchmarking: surgical volume and operative mortality in Veterans Administration Medical Centers. *Best Pract Benchmarking Healthc.* 1996;1:34-42. [PMID: 9192589]
10. Elixhauser A, Steiner C, Fraser I. Volume thresholds and hospital characteristics in the United States. *Health Aff (Millwood).* 2003;22:167-77. [PMID: 12674419]
11. Localio AR, Berlin JA, Ten Have TR, Kimmel SE. Adjustments for center in multicenter studies: an overview. *Ann Intern Med.* 2001;135:112-23. [PMID: 11453711]
12. Williamson JM, Datta S, Glen AS. Marginal analyses of clustered data when cluster size is informative. *Biometrics.* 2003;59:36-42. [PMID: 12762439]
13. Laine C, Markson LE, McKee LJ, Hauck WW, Fanning TR, Turner BJ. The relationship of clinic experience with advanced HIV and survival of women with AIDS. *AIDS.* 1998;12:417-24. [PMID: 9520172]
14. Kitahata MM, Van Rompaey SE, Dillingham PW, Koepsell TD, Deyo RA, Dodge W, et al. Primary care delivery is associated with greater physician experience and improved survival among persons with AIDS. *J Gen Intern Med.* 2003;18:95-103. [PMID: 12542583]

© 2003 American College of Physicians