Provider volume and clinical outcomes in surgery: Issues and implications

by
Clara N. Lee, MD, MPP,
and
John M. Daly, MD, FACS,
New York, NY

Concern about quality in health care has undergone dramatic resurgence in recent years. Unfortunately, the availability of objective data on the quality of care delivered by individual surgeons and hospitals has remained limited. For this reason, patients, insurers, corporate purchasers, and policymakers are looking increasingly to provider volume as a surrogate indicator for quality of care.

A growing body of literature finds a positive association between hospital and surgeon volume and clinical outcomes in surgery. The magnitude and the nature of this association appear to be highly variable, however, depending on the procedure and on the study design. The methodological rigor of investigations on volume and outcomes is also highly variable. For surgeons to take part in this discussion and the policy debate that arises from it, we need to understand the conceptual issues underlying the debate and to be able to critically assess the literature.

This article explains the conceptual issues that form the basis of the discussion about provider volume and clinical outcomes and proposes a method to evaluate the literature. It reviews some of the findings in the literature and then describes the policy implications of those findings.

Conceptual issues

Provider volume itself is not the equivalent of health care quality. Rather, it serves as a proxy for quality, because quality itself is difficult to define and difficult to measure. Many initiatives under way hold the promise of eventually improving the measurement of health care quality.[1] In the meantime, however, clinical volume will continue to be used as an indicator of quality and a predictor of outcome.

Two hypotheses have been proposed to explain the association between volume and outcomes. The "practice makes perfect" hypothesis is based on the notion that repetition improves a surgeon's or hospital's ability to perform a procedure. The "selective referral" hypothesis stems from the premise that those providers who have good outcomes receive more referrals as a result, and therefore have higher volumes. The practice makes perfect hypothesis has intuitive appeal but begs several questions. Is there a limit to how much a surgeon or hospital can improve with increasing volume? Is there a threshold of experience that a provider must obtain before seeing any improvement in quality? Do all surgical procedures improve with higher volumes? One would expect, given the practice makes perfect hypothesis, that for simple procedures, practice may not be necessary, whereas for more complex operations, practice would be important; that is to say, a strong relationship between volume and outcomes would exist for esophagectomy but not for breast lumpectomy.

For the selective referral hypothesis to be true, two conditions must be met. First, patients and referring physicians must know which surgeons and hospitals have better outcomes. Second, they must be willing and able to act upon such knowledge. With a few exceptions, such as the public reporting of mortality statistics for cardiac surgery

in certain states, outcome data for individual surgeons and hospitals are not, in fact, available to the public. Even when such data are available, referring physicians still do not base their referrals on that information,[2,3] and patients do not necessarily use that data when choosing providers.[4,5] The selective referral hypothesis also requires that patients have their choice of surgeons and hospitals, but many patients do not for financial, insurance, logistical, or other reasons.[6]

An alternative way to think about the relationship between volume and outcomes is to ask just what high-volume providers do differently than low-volume providers. Do they take specific actions that others do not? If so, which of those actions affect outcome? Health services researchers would call such specific actions "processes of care." An example of a process of care is a specific maneuver that a vascular surgeon does during a carotid endarterectomy, or the cross-clamp time during coronary artery bypass grafting. The more we learn about the differences in specific processes of care and their effects on patient outcomes, the more we could potentially improve the outcomes of low-volume providers.

Critically assessing the literature

The literature on volume and outcomes in surgery is growing rapidly, and the quality of the literature is highly variable. The ideal study should have findings that could be generalized beyond the sample of the study and could provide some insight regarding the magnitude and nature of the relationship between volume and outcome. Thus, a surgeon reading these studies should examine several factors, including the type and size of the sample being studied, the use of risk adjustment, the quality and type of data used, and the measurement of clinical processes of care.

- *Sample*. For a study to be truly generalizable, it must examine a population-based sample of patients and all surgeons/hospitals available to that population-based sample. An example of a population-based sample would be all patients undergoing coronary artery bypass graft in New York State in a given year. An example of a nonpopulation-based sample would be all carotid endarterectomy patients seen at two hospitals. Such a study might still provide valuable information, but its conclusions could not be considered those from which generalizations could be drawn.

A study should also include enough variability among the providers being compared. For example, one study of pancreaticoduodenectomy in Maryland included 15 hospitals, but only one institution qualified as "high-volume." Thus, the ability to generalize its findings to all high-volume hospitals is limited.

- *Risk adjustment*. One of the most important aspects of any outcome assessment is risk adjustment. Risk adjustment is the process of accounting for differences in patients' comorbidities and severity of disease. It enables us to determine whether any differences in outcome are due to differences in the care given or merely to differences in the patients. Some have argued that higher-volume surgeons may have better outcomes merely because they operate on less sick patients, rather than because of anything that they do better.

The type of data used for risk adjustment is also important. Most studies use administrative data, such as diagnosis-related codes, for information about patient comorbidities. This method is fraught with inaccuracies.[7] Very few studies use the far more accurate source of data on comorbidities and severity—namely, clinical information abstracted from patient charts.

- *Data type*. The source of data, in general, is an important factor in determining a study's validity. The most commonly used sources of outcome data are administrative data banks. Such databases are designed for billing purposes, usually only contain information on mortality, and are notoriously limited in their accuracy.[7,8] A better source of data would be a repository of clinical, rather than administrative, information. For example, Medicare's surveillance, epidemiology, and end results (SEER) database was specifically designed to collect clinical details about each patient in the database, rather than merely discharge diagnoses. Finally, the ideal source of data on clinical outcomes would be the actual clinical record. Specific pieces of information about the outcome of a case, such as death, stroke, myocardial infection, and so on, could be obtained.
- *Processes of care*. A few studies measure how high-volume providers differ from low-volume providers in terms of what they actually do. For example, some of the studies of oncologic surgery

adjust for the use of adjuvant therapy or for the type of surgical resection.[9,10] Another example is a study of total hip replacement that adjusted for the type of prosthesis used and for the use of perioperative antibiotics.[11] Such data on what high-volume surgeons actually do differently from low-volume surgeons would have major implications for management algorithms to improve quality.

Literature findings

The literature on volume and outcomes in surgery has become quite extensive, covering many types of surgery, including pancreatic, esophageal, lung, colorectal, breast, cardiac, pediatric cardiac, joint replacement, carotid, and abdominal aortic aneurysm surgery. The reader is referred elsewhere for more exhaustive reviews of the literature.[12,13] Here, we review some of the most significant findings.

The majority of studies find a positive association between higher volume and better outcome.[12] This is true for studies of surgeon volume and for studies of hospital volume.

In general, the more complex surgical procedures show the strongest association between provider volume and outcomes (see Tables 1 and 2, this page). Among all of the procedures studied, the strongest association exists for surgery of the pancreas and esophagus and in pediatric cardiac surgery. One study of outcomes after pancreaticoduodenectomy reported a mortality difference of 7 percent between high-volume and low-volume surgeons. It also found a mortality difference of 14 percent for high-volume hospitals compared to low-volume hospitals.[14] Studies of outcomes after esophagectomy reported a mortality difference between low-volume versus high-volume hospitals ranging from 10 to 13 percent.[15-17] Differences in mortality were more modest for CABG, carotid endarterectomy, lung resection, and colorectal surgery.

A somewhat surprising finding is that even for some relatively simple operations, clinical outcomes improve with higher volumes. One study found that the five-year survival rate for breast cancer patients undergoing mastectomy or lumpectomy was higher at high-volume institutions.[10] Similarly, another study found that survival after mastectomy or lumpectomy was higher in the hands of high-volume surgeons.[18] Because these operations are not particularly complex, the authors of these studies attributed the differences to better coordination of care and more appropriate use of adjuvant therapy by high-volume sur-

Table 1

Difference in mortality between high- and low-volume hospitals

Procedure	Mortality difference (%)*
Pancreatectomy	3-17
Esophagectomy	11-14
Peds cardiac	2-15
Lung resection	3-4
Colorectal	1.7-1.9
CABG	1-4
Carotid endarterectomy	.5-2

*Reported differences in mortality between high-volume and low-volume hospitals, based on a systematic review of the literature (Halm and others, 2001[12]).

Table 2

Difference in mortality between high- and low-volume surgeons

Procedure	Mortality difference (%)*
Pancreatectomy	6-7
Lung resection	0
Colorectal	1.86-1.90
Abdominal aortic aneurysm	5-12
Carotid endarterectomy	1-4

*Reported differences in mortality between high-volume and low-volume hospitals, based on a systematic review of the literature (Halm and others, 2001[12]).

geons and hospitals. Thus, provider volume appears to matter not just for technical reasons.

It is important to remember that although higher patient volumes appear to be associated with better outcomes, this is true only on average.[19] A large amount of variation in outcomes still exists among individual surgeons and hospitals. Thus, some high-volume surgeons have poor outcomes, and some low-volume surgeons have very good outcomes.[20,21] Policies based solely on provider volume, such as steering patients toward higher-volume hospitals, therefore, might improve outcomes on average but not in every case.

One reason for the variability among individual surgeons is that some high-volume surgeons operate at low- or intermediate-volume institutions, and some low-volume surgeons operate at intermediate- or high-volume institutions. Although we know that both hospital and surgeon volume affect outcome, we do not know what their relative contributions are or how they interact, because very few studies have evaluated both surgeon volume and hospital volume together.[12] Policies based solely on surgeon volume without taking into account hospital volume (and vice versa), therefore, will not always have their intended effects.

One study of coronary artery bypass in New York State evaluated both surgeon volume and hospital volume.[22] It found that high-volume surgeons at high-volume hospitals had a risk-adjusted mortality of 2.18 percent, whereas low-volume surgeons at low-volume hospitals had a risk-adjusted mortality of 14 percent. High-volume surgeons had lower mortality rates than low-volume surgeons, regardless of hospital volume.

A recent study of common cancer procedures also evaluated the effects of both surgeon volume and hospital volume.[23] It found that for colectomy, lung lobectomy, and gastrectomy, the lowest risk-adjusted mortality rates were for high-volume surgeons at high-volume hospitals, and the highest risk-adjusted mortality rates were for low-volume surgeons at low-volume hospitals. Patients operated on by high-volume surgeons at high-volume hospitals usually had lower mortality rates than patients operated on by low-volume surgeons or in low-volume hospitals, or both.

Although it does appear that higher volumes are associated with better outcomes, a number of questions remain regarding the specific nature of that association. Does the association between volume and outcome diminish over time, once a procedure has existed for a longer time and has become more widespread? Is there a threshold effect—a certain volume beyond which there is no more improvement in outcomes with increasing volume? Is there a second threshold effect—a volume beyond which outcomes are worse?

How do surgeons' outcomes change as they perform a given procedure more times? What is the shape of the learning curve? In particular, what is the shape of the learning curve for a newer surgeon learning a standard procedure or for a more experienced surgeon learning a new procedure? Is there a constant procedure volume that a surgeon must maintain to remain competent at performing that procedure? Is the cumulative lifetime volume more important? Does the surgeon's volume with other, similar procedures contribute to the outcome?

Another important limitation of the literature is its focus on mortality as the outcome. Only a few studies measured complications as the outcome of interest. Even fewer studies measured such endpoints as long-term survival, functional status, recurrence of disease, or quality of life.

Finally, although the literature strongly supports the notion that higher volumes are associated with better outcomes, particularly for complex surgery, it sheds very little light on why or how. One of the studies that measured specific "processes of care" found that much of the difference in survival from myocardial infarction at

Dr. Lee *is associate instructor, Joan and Sanford Weill Medical College, Cornell University, and assistant surgeon, New York Presbyterian Hospital, New York, NY.*

higher-volume institutions was due to greater use of proven-effective medications (aspirin and beta-blockers).[24] To truly explain the mechanisms underlying the volume-outcome association, future research would need to measure specific details about what higher-volume providers do differently than lower-volume providers.

Policies

A number of policies have been advocated based on the association between surgeon and hospital volumes and outcomes. One approach would be to publicize data on surgeons' and hospitals' volumes for specific procedures. This could take the form of public databases, Web sites, "report cards," or employer-sponsored information for employees. The success of this type of approach would depend on how well the public could be educated about the complexities of the relationship between provider volume and outcomes. In addition, it would depend on patients' ability to use this information when choosing a hospital or surgeon—an ability that has been shown to be limited.[25,6]

Some large health care purchasers, such as business coalitions and managed care organizations, already are seeking to selectively contract with higher-volume institutions and to encourage referrals to higher-volume surgeons. The Pacific Business Group on Health, a large health care purchasing coalition, requires its health plans to steer patients with certain conditions toward higher-volume institutions.

Another policy approach would be for state governments to regionalize the provision of certain services, thereby boosting volumes at a few institutions. Such policies have been implemented in the past for trauma and for newborn intensive care. In addition, some states have certificate of need programs that limit the number of hospitals allowed to perform certain procedures, such as cardiac surgery and organ transplantation.

The effect of a large movement of patients away from lower-volume providers toward higher-volume providers is not well understood and could include some unintended consequences. As higher-volume institutions developed greater market share, they would have greater ability to raise their fees. Thus, regulation of fees might need to be implemented concurrently. Hospitals might have a new, medically inappropriate incentive to admit or to operate on a patient in order to boost volumes. Finally, the use of volume criteria could make it more difficult for new surgeons and new institutions to break into the market.[25] Special provisions for new surgeons or for surgeons-in-training might be necessary.

A final approach for improving outcomes would be to implement quality improvement programs. Research that identifies specific processes of care associated with higher volumes could be used to improve the performance of lower-volume surgeons through retraining and education programs. In order for this to be possible, future studies would need to measure specific processes of care. Databases containing such specific data and based on population-based samples would be extremely helpful to this end.

Conclusion

Overall, higher surgeon and hospital volumes appear to be associated with better clinical outcomes, particularly for less common and complex operations. The literature on the subject is quite extensive but has a number of methodological limitations. It leaves unanswered many questions regarding the specific nature and magnitude of the association between volume and outcomes.

Most of the policies that have been proposed seek to shift patients away from lower-volume providers toward higher-volume providers, either through public education efforts, selective referral and contracting, or government regionalization of care. An alternative approach would be quality improvement programs in which lower-volume

Dr. Daly *is Lewis Atterbury Stimson Professor and chair, department of surgery, Joan and Sanford Weill Medical College, Cornell University, and surgeon-in-chief, New York Presbyterian Hospital (Cornell), New York, NY.*

providers could learn from the practices of higher-volume surgeons and hospitals. Rather than using volume as a proxy for quality, such an approach would seek to improve the actual quality of care being rendered. Ω

References

1. National Committee for Quality Assurance: *A Road Map for Information Systems: Evolving Systems to Support Performance Measurement*. Washington, DC: National Committee for Quality Assurance, 1997.
2. Schneider EC, Epstein AM: Influence of cardiac surgery performance reports on referral practices and access to care. A survey of cardiovascular specialists. *N Engl J Med*, 335(4):251-256, 1996.
3. Hannan EL, Stone CC, Biddle TL, et al: Public release of cardiac surgery outcomes data in New York: What do New York State cardiologists think of it? *Am Heart J*, 143(6):1120-1128, 1997.
4. Chassin MR, Hannan EL, DeBuono BA: Benefits and hazards of reporting medical outcomes publicly. *N Engl J Med*, 334(6):394-398, 1996.
5. Schneider EC, Epstein AM: Use of public performance reports: A survey of patients undergoing cardiac surgery. *JAMA*, 279(20):1638-1642, 1998.
6. Marshall MN, Shekelle PG, Leatherman S, et al: The public release of performance data: What do we expect to gain? A review of the evidence. *JAMA*, 283(14):1866-1874, 2000.
7. Donaldson MS: *Measuring the Quality of Health Care*. Washington, DC: National Academy Press, 1999.
8. Hannan EL, Racz MJ, Jollis JG, et al: Using Medicare claims data to assess provider quality for CABG surgery: Does it work well enough? *Hlth Serv Res*, 31(6):659-678, 1997.
9. Porter GA, Soskolne CL, Yakimets WW, et al: Surgeon-related factors and outcomes in rectal cancer. *Ann Surg*, 227(2):157-167, 1998.
10. Roohan PJ, Bickell NA, Baptiste MS, et al: Hospital volume differences and five-year survival from breast cancer. *Am J Pub Hlth*, 88(3):454-457, 1998.
11. Espehaug B, Havelin LI, Engesaeter LB, et al: The effect of hospital-type and operating volume on the survival of hip replacements. A review of 39,505 primary total hip replacements reported to the Norwegian Arthroplasty Register, 1988-1996. *Acta Ortho Scand*, 70:12-18, 1990.
12. Halm EA, Lee C, Chassin MR: Is volume related to outcome in health care? A systematic review and methodologic critique of the literature. *Ann Int Med*, 2001; in press.
13. Hannan EL: *Provider Volume-Patient Outcome Relationships in the Provision of Medical Care: An Update*. Rockville, MD: Agency for Health Research and Quality, 2001.
14. Lieberman MD, Kilburn H, Lindsey M, et al: Relations of perioperative deaths to hospital volume among patients undergoing pancreatic resection for malignancy. *Ann Surg*, 222(5):638-645, 1995.
15. Gordon TA, Bowman HM, Bass EB, et al: Complex gastrointestinal surgery: Impact of provider experience on clinical and economic outcomes. *J Am Coll Surg*, 189(1):46-56, 1999.
16. Begg CB, Cramer LD, Hoskins WJ, et al: Impact of hospital volume on operative mortality for major cancer surgery. *JAMA*, 280(20):1747-1751, 1998.
17. Patti M, Corvera CU, Glasgow RE, et al: A hospital's annual rate of esophagectomy influences the operative mortality rate. *J Gastro Surg*, 2:186-192, 1998.
18. Sainsbury R, Haward B, Rider L, et al: Influence of clinician workload and patterns of treatment on survival from breast cancer. *Lancet*, 345(8960): 1265-1270, 1995.
19. Hannan EL, O'Donnell JF, Kilburn H, et al: Investigation of the relationship between volume and mortality for surgical procedures performed in New York State hospitals. *JAMA*, 262(4):503-510, 1989.
20. Chassin MR, Park RE, Lohr KN, et al: Differences among hospitals in Medicare patient mortality. *Health Serv Res*, 24(1):1-31, 1989.
21. Park RE, Borrk RH, Kosecoff J, et al: Explaining variations in hospital death rates: Randomness, severity of illness, quality of care. *JAMA*, 264(4):484-490, 1990.
22. Hannan EL, Kilburn H, Bernard H, et al: Coronary artery bypass surgery: The relationship between inhospital mortality rate and surgical volume after controlling for clinical risk factors. *Med Care*, 29:1094-1107, 1991.
23. Hannan E, Radzyner M, Rubin D, et al: The influence of hospital and surgeon volume on in-hospital mortality for colectomy, gastrectomy, and lung lobectomy in patients with cancer. *Surgery*, 131(1):6-15, 2002.
24. Thiemann DR, Coresh J, Oetgen WJ, et al: The association between hospital volume and survival after acute myocardial infarction in elderly patients. *N Engl J Med*, 340(21):1640-1648, 1999.
25. Hewitt M, Petitti D: *Interpreting the Volume-Outcome Relationship in the Context of Cancer Care*. Bethesda, MD: Institute of Medicine, National Academy Press, 2001.