Presidential Address

Society of Surgical Oncology Presidential Address: Volume, Outcome, and Surgical Specialization

John M. Daly, MD

The opportunity to address you today is a singular privilege, one that I do not take lightly. Many of you are mentors, colleagues, trainees, friends, and family who have shaped this talk through your personal interactions, your writings, and your support. For all you have done, I am eternally grateful.

Presidential talks vary in theme from personal messages to lifelong research projects, from philosophy to biologic topics. I have chosen to talk on volume, outcome, and specialty training as a means to expand from my personal observations to public policy issues and from individual education to a commitment toward societal good.

Just as “a journey of a thousand miles starts with a single step,” so too, a large clinical database starts with a single patient. In 1997, a young woman was found clinically to have a large pelvic and abdominal mass. She and her family were frightened. They had no specific knowledge of a hospital’s outcome results with this problem; they sought help on the basis of institutional reputation and physician availability. After preoperative testing, the woman underwent a total abdominal hysterectomy, a bilateral oophorectomy, an appendectomy, and an abdominal staging procedure. An adenocarcinoid tumor of the appendix was the primary culprit, and this situation represented stage 4 disease. Subsequently, a catheter was placed for intraperitoneal chemotherapy, followed by intravenous therapy, although definitive outcome data were not available for this tumor type and stage. A major pulmonary embolus occurred after hospital discharge, requiring readmission.

During this time, knowing the prognosis, the patient and her family sought additional treatment options. Eight months after the first procedure, a second-look staging operation and hyperthermic intraperitoneal infusion were done, followed by additional intraperitoneal and intravenous chemotherapy for a total treatment period of 2 years. Institutional volume in this instance was a single case, but surgeon experience was extensive. After the patient’s chemotherapy was completed, she went on to receive experimental vaccine therapy, knowing again that results for this tumor type were unclear.

What does this individual patient teach us? First, institutional and surgeon reputations, not outcomes data, play major roles in an individual’s treatment selection. Second, complications do occur after hospital discharge and may not always be recorded in certain databases. Third, the benefits of certain treatment options relate to averages and are not specific to individuals. Finally, hospital and surgeon volumes, along with specialty training, may influence patient outcome, but the processes of care that are most important are not always clear.

Concern over quality in health care has seen dramatic resurgence in recent years. Underuse, overuse, and misuse of hospitalized patient care continues to occur. Gawande et al., in a study of adverse events in Colorado and Utah, noted that 54% of surgical adverse events were preventable, and 6% of these adverse events resulted in death. 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 look increasingly to provider volume as a surrogate indicator for quality of care (Table 1). In May 2000, for example, the Institute of Medicine convened a workshop to discuss hospital and surgeon volumes and clinical outcomes, along with recommending public policies for selective patient referral to institutions.

A growing body of literature finds a positive association between hospital and surgeon volumes and clinical outcomes in surgery (Table 2). The magnitude and the
nature of this association appear to be highly variable, however, depending on the procedure and the study design. For example, the association is much clearer for complex operations such as pancreaticoduodenectomy, esophagectomy, and hepatic resection. For surgeons to take part in this discussion and the policy issues that arise from it, we need to thoroughly understand the conceptual problems underlying the debate.

This address focuses on the issues that form the basis of the discussion about provider volumes and clinical outcomes. It reviews some published reports and then describes the policy implications of those findings.4–8

**Conceptual Issues**

There are a number of factors that affect treatment processes and outcomes. Patient education, income, and comorbidities, insurer access, coordination of services, and geography all play important roles.9–12

Provider volume itself is not the equivalent of health care quality. Rather, it serves as a proxy for quality, since quality itself is difficult to define and difficult to measure. Many initiatives under way hold promise to eventually improve the measurement of health care quality. In the meantime, however, clinical volume will continue to be used as a surrogate indicator of quality and a predictor of outcome. Halm et al. reviewed 88 published reports and found that 77% of these studies showed a significant association between volume and outcome.7 However, we do not completely understand what specific factors explain outcome differences.

Two hypotheses have been proposed to explain the association between volumes and outcomes. The practice-makes-perfect hypothesis is the notion that repetition improves a surgeon’s or hospital’s ability to perform a procedure. The selective referral hypothesis is the notion that those providers who have good outcomes receive more referrals as a result and therefore have higher volumes.13

The practice-makes-perfect hypothesis has intuitive appeal but also begs several questions. Is there a limit to how much a surgeon or hospital improves with increasing volume? Is there a threshold of experience that a provider must obtain before seeing any improvements 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—for example, a strong relationship between volumes 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 on individual surgeons and hospitals are not in fact available to the public.14,15 Even when such data are available, referring physicians still do not base their referrals on that data, and patients do not necessarily use that data in choosing their providers. The selective referral hypothesis also requires that patients have a choice among surgeons and hospitals, but many patients do not, for financial, insurance, logistical, and other reasons.

An alternative way to think about the relationship between volumes 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 make a difference in outcomes? Health services researchers would call such specific actions processes of care, as shown in Fig. 1.7

Please see the print version for this Figure.

An example of a process of care is a specific maneuver that a surgeon makes 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 all institutions and surgeons.

The literature on volumes and outcomes in surgery is growing rapidly, but the quality of the literature is highly variable. The ideal study 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 (Table 3).

---

**TABLE 1.** Quality in health care

<table>
<thead>
<tr>
<th>Patients</th>
<th>Insurers</th>
<th>Policy makers</th>
<th>Government purchasers</th>
<th>Corporate purchasers</th>
</tr>
</thead>
</table>

---

**TABLE 2.** Mortality

<table>
<thead>
<tr>
<th>Procedure</th>
<th>N</th>
<th>1–5</th>
<th>6–10</th>
<th>11+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pancreatectomy</td>
<td>19,000</td>
<td>13%</td>
<td>8%</td>
<td>6%</td>
</tr>
<tr>
<td>Esophagectomy</td>
<td>6,782</td>
<td>17%</td>
<td>4%</td>
<td>3%</td>
</tr>
<tr>
<td>Hepatic resection</td>
<td>126,395</td>
<td>5.4%</td>
<td>3.5%</td>
<td>1.7%</td>
</tr>
</tbody>
</table>

Reprinted from *JAMA*, 1998;280:1747–51, Copyright © (1998), American Medical Association. All rights reserved.
Sample

For a study to be truly valuable, it must examine a population-based sample of patients and all surgeons and hospitals available to that population-based sample. An example of a population-based study would be all patients undergoing coronary artery bypass graft operations in New York State in a given year. An example of a non-population-based sample would be all carotid endarterectomy patients seen at two hospitals.

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

Risk Adjustment

One of the most important aspects of any outcome assessment is risk adjustment. 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 (e.g., DRG codes) for information about patient comorbidities. This method is fraught with inaccuracies. Very few studies utilize far more accurate sources of data on comorbidities and severity, namely, clinical information abstracted from patient charts. Specific pieces of information about the outcome of a case, such as stroke, myocardial infarction, and pneumonia, could then be obtained.

Processes of Care

A few studies measure how providers differ in terms of what they actually do. For example, some studies of cancer surgery adjust for the use of adjuvant therapy or for the type of surgical resection. While results of colon resections may not differ much between surgeons or institutions, the results of rectal resections have shown substantial differences. Such data on what high-volume surgeons actually do differently from low-volume surgeons would have major implications for management algorithms to improve quality.

TABLE 3. Outcome studies

| - Population-based patient sample |
| - All surgeons/hospitals available to population |
| - Study should provide variability in outcomes |
| - Risk adjustment: |
| - Comorbidities |
| - Severity of disease |
| - Processes of care |

Literature Findings

The literature on volumes and outcomes in surgery has become quite extensive, covering many types of surgery, including pancreatic, esophageal, lung, colorectal, breast, cardiac, carotid, and abdominal aortic aneurysm surgery.

The majority of studies find a positive association between higher volumes and better outcomes. This is true for studies of surgeon volume and for hospital volume.

In general, the more complex surgical procedures show the strongest association between provider volume and outcomes. Among all oncologic procedures studied, the strongest association exists for surgery of the pancreas and esophagus, as shown in Fig. 2. Studies of outcomes after esophagectomy reported a mortality difference between hospitals ranging from 10% to 13%, as shown in Fig. 3. Studies of outcomes after pancreaticoduodenectomy revealed a mortality difference between high-volume and low-volume surgeons of 7% and a mortality difference of 10% for hospitals. These differences have been noted with use of large state-wide and national databases, as shown in this composite graph. In a report by Dudley et al., potentially avoidable deaths in California were noted in 5% to 7% of cases of esophageal and pancreatic resections in which the patients were admitted to low-volume hospitals. In another study, by Birkmeyer and colleagues, survival was improved at 3 years among patients undergoing pancreatectomy at high-volume centers. Differences in mortality were more modest for lung resection and colorectal surgery.

A somewhat surprising finding is that even for some relatively simple operations, choice of procedure is dif-

FIG. 2. With complex surgical procedures such as esophagectomy and pancreatic resection, increasing hospital volume leads to lower mortality.
Different between high- and low-volume centers, with lumpectomy more commonly performed by surgeons with higher volumes. In addition, longer-term clinical outcomes are improved with higher volumes. One study found that the 5-year survival rate for breast cancer patients undergoing mastectomy or lumpectomy was better at high-volume institutions.\textsuperscript{28–30} Similarly, another study found that survival after mastectomy or lumpectomy was higher for patients of high-volume surgeons. Since 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 practitioners at high-volume hospitals. Use of multidisciplinary, disease-focused clinics or tumor boards may provide the critical difference. 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, as shown in Fig. 5.\textsuperscript{31} 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. Policies based solely on provider volume, such as steering patients toward certain hospitals, therefore, might improve outcomes on average but not in every case.

As shown in the results from Hannan et al., hospital and surgeon volumes tend to correlate with one another.\textsuperscript{31} However, one of the reasons for the variability among individual surgeons is that some high-volume surgeons operate at low-volume institutions, and some low-volume surgeons operate at high-volume institutions. Although we know that both hospital and surgeon volumes 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. Policies based solely on surgeon volume, without taking into account hospital volume (and vice versa), will therefore not always have their intended effects.

A recent study of common cancer procedures evaluated the effects of both surgeon volume and hospital volume (Table 4). Hannan and colleagues\textsuperscript{8} found that for
colocemy 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 in other categories showed intermediate results.\textsuperscript{17,22,31}

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 such that beyond a certain number of patients or procedures there is no more improvement in outcomes with increasing volume? And is there a second threshold, a number beyond which outcomes are worse? For example, Cox et al. noted that surgeons at the Moffit Cancer Center had a 90% success rate in identifying a sentinel lymph node after 22 cases, a rate which remained steady.\textsuperscript{32} Their database also indicated that in other surgeons hands, an average of six cases per month resulted in a 97% success rate.

How do surgeons’ outcomes change as they perform a given procedure more times? That is, 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 rate that a surgeon must maintain in order to remain capable of performing that procedure? Is the cumulative lifetime number more important? Does the surgeon’s volume with other, similar procedures contribute to the outcome?

An important limitation of knowledge is the current focus on mortality as the outcome. Few studies measure complications as the outcome of interest. Even fewer studies measure such endpoints as long-term survival, functional status, recurrence of disease, and 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. Specialty training is emerging as having a major independent role in clinical outcomes.\textsuperscript{33} This has been shown for patients undergoing carotid endarterectomy by vascular surgeons, general surgeons, and neurosurgeons. It has been suggested also by results for patients undergoing treatment for intermediate-stage ovarian cancer, as shown in Table 5. Using a large state database as shown in Fig. 6, we noted a statistically significant reduction in mortality among patients undergoing gastrectomy and colectomy performed by members of the Society of Surgical Oncology and the American Society of Colorectal Surgeons, respectively; this reduction was independent of hospital and surgeon volume. \textit{You do make a difference.} To truly explain the mechanisms underlying the volume-outcome association, future research would need to measure specific details about what you do differently than other providers.

### Policies

A number of policies have been advocated on the basis of the association between surgeon and hospital volumes and outcomes (Table 6).\textsuperscript{34} One approach would be to publicize data on surgeons’ and hospitals’ volumes for specific procedures.\textsuperscript{35} This could take the form of public databases, newspapers, 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 volumes and outcomes. In addition, it would depend on patients’ abilities to make use of this information when choosing a hospital or surgeon, an ability that has been shown to be limited. For example, in one California study, 25% of patients traveled further to reach a low-volume institution, bypassing a high-volume center in the process.\textsuperscript{35}

Some large purchasers of health care, such as business coalitions and managed care organizations, are already 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

<table>
<thead>
<tr>
<th>Stage</th>
<th>Gyn Onc</th>
<th>OBG</th>
<th>Gen Surg</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>89%</td>
<td>90%</td>
<td>88%</td>
</tr>
<tr>
<td>2</td>
<td>63%</td>
<td>61%</td>
<td>47%</td>
</tr>
<tr>
<td>3</td>
<td>25%</td>
<td>29%</td>
<td>17%</td>
</tr>
<tr>
<td>4</td>
<td>10%</td>
<td>17%</td>
<td>11%</td>
</tr>
</tbody>
</table>

\textit{Table 5. Ovarian cancer: survival by specialty}
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. Restrictions to care, if they occur, should initially be in metropolitan areas where alternatives exist and should be focused on procedures for which the greatest differences in outcome occur. However, certificate-of-need medical programs are decreasing in many regions. An interesting note is that the outcome of coronary bypass graft was not associated with volume once certain standards were met. This was not due to a greater proportion of operations performed by high-volume surgeons, but the gap in outcome results narrowed between low- and high-volume providers. This is the ultimate objective.

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. Overuse of surgical interventions might occur on a grand scale. Hospitals might have a new, medically inappropriate incentive to admit or operate on a patient, that is, 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. Special provisions for new surgeons or for surgeons in training might be necessary.

Another approach for improving outcomes would be to fully implement quality improvement programs. Better measures of quality are needed that are truly risk-adjusted. 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.

In Conclusion

Higher surgeon and hospital volumes appear to be associated with better clinical outcomes, particularly for less common and complex operations. The research 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 certain institutions and surgeons toward higher-volume providers, through public education efforts, selective referral and contracting, or government regionalization of care. An alternative approach would be quality improvement programs in which providers could learn from the practices of other surgeons and hospitals to improve clinical outcomes. Rather than using volume as a proxy for quality, such an approach would seek to improve the actual quality of care being rendered. There is good evidence that physicians do change their practice patterns if treatment guidelines are local, not national, if you as opinion leaders promulgate these guidelines, and if there is wide local participation and feedback of results. The Society of Surgical Oncology exists as a forum to discuss new ideas and therapies for you, the experts in the field. It is our mantle of responsibility to incorporate these into guidelines, to share and promote them among our colleagues, and to improve the quality of oncologic care for all patients.

Some of you may be wondering about the conclusion of the story regarding the young woman in 1997. She is still unaware of clear outcomes data regarding her therapies, which became bicoastal and experimental. Instead, she has developed an unwavering faith in her nurses, her doctors, her family, and her God. She is learning to accept ambiguity regarding tests and the deaths of her

![Image](image_url)
friends whom she came to know during her initial chemotherapy treatments. She is truly an N = 1; she is my wife of 31 years and this is her story.

Thank you for the privilege of being your president this past year!

REFERENCES