The Volume–Outcome Debate Revisited
Samuel R.G. Finlayson, MD, MPH
From the Department of Surgery, Dartmouth Medical School, Hanover, NH
and the Veteran’s Administration Outcomes Group, White River Junction, VT

ABSTRACT
Multiple studies support the intuitive association between higher provider procedure volume and better clinical outcomes. Healthcare purchasers and payers have been seeking ways to direct patients to high-volume providers to improve the quality of care received and to avoid costs associated with higher surgical morbidity. Volume-based referral has faced resistance from providers who are concerned that the use of volume instead of more direct measures of surgical quality will result in unfair discrimination. On close examination, volume-based referral policies also appear to be more congruent with payers’ interests than the interests of individual patients and providers. Furthermore, a policy of volume-based referral does not address surgical quality directly, is applicable to only a very small segment of surgical care, and is logistically problematic. However, in the absence of viable alternative measures of surgical quality, imperfect proxies such as volume will likely continue to be a significant part of the national dialogue surrounding surgical quality.

In recent years, many studies have shown that higher procedure volume is associated with better clinical outcomes. The functional corollary of this observation is that high-risk operations should be referred to high-volume providers in order to improve the safety and quality of surgical care (also known as regionalization). Attempts to quantify the benefit of volume-based referral have resulted in very large estimates of potential lives saved.1-3 Payers’ enthusiasm for such a strategy is also driven by the potential cost savings that would result if these procedures could be performed at lower rates of morbidity and mortality. Advocates of volume-based referral would argue that passive acquiescence to the status quo systematically harms patients while simultaneously wasting resources in a healthcare system that is already economically strained. Although few volume-based initiatives have been systematically applied or compulsory, some healthcare purchasers (eg, Leapfrog Group) have strongly advocated volume-based referral, and a growing number of measures designed to restrict specialty surgery (eg, transplantation and bariatric) to accredited high-volume providers are emerging.

Opponents of a policy of volume-based referral have been equally vocal. Foremost among the arguments against regionalization is the contention that individual providers should be judged by their own clinical results, not a generalized statistical inference about volume and mortality. Volume is seen as a “blunt instrument” that lacks the precision to determine with any real degree of certainty whether an individual provider performs at a high standard of quality. Insofar as volume is a proxy for quality, not quality itself, many providers see volume-based referral as unfair discrimination.

These arguments for and against regionalization are based on valid and reasonable principles. Neither side of the debate can claim sole occupation of the moral high ground. Virtually everyone would agree that patient safety and well-being, economic responsibility, and fairness are all interests worth defending. In this article, I will discuss the volume–outcome association and the shortcomings of regionalization as a primary lever to improve surgical quality, and in the process present a different view of the debate—a view based on consideration of the distinct perspectives of patients, providers, and payers. This view is intended to deepen the discourse and, hopefully, help bring us closer to a system of surgical care that appropriately considers the interests of all stakeholders in the health care system—patients, providers, and payers.

Volume–Outcome: Is It Real?

In 1979, Luft et al published a landmark article in the New England Journal of Medicine that observed the relationship between surgical volume and mortality and framed the argument for regionalization.4 Since then, numerous studies have documented this relationship for many different surgical procedures. The sizable body of literature on the subject is concisely summarized in review articles authored by Dudley et al in 20003 and Halm et al in 20025 (Table 1). Studies of the volume–outcome relationship have been not only numerous but also consistent, with 123 out of 128 studies reviewed by Dudley et al showing lower rates of mortality associated with higher volume for over 40 different procedures.

The volume–outcome observation has long prompted skepticism. Initial doubts included the contention that...
Volume–Outcome as a Quality Lever: Whose Interests Are Served?

Although volume–based referral has gained increasing acceptance over time, the idea still creates some degree of uneasiness among providers. While much of this may be related to surgeons’ subjective fears of losing market share, a more objective view does yield legitimate concerns about the use of provider volume as a primary determinant of surgical quality. Some of these concerns relate to how volume-based referral addresses the individual interests of the three main stakeholders in surgical care: payers, providers, and patients.

Volume-Based Referral and Payers

Although some patient advocates and high volume providers have been vocal backers of volume-based referral, payers have delivered the greatest impetus in favor of these measures. Payers seek not only improvement in the quality of care provided to their customers, but also stand to gain the most economically from lower morbidity and mortality. An obvious explanation for the central role payers have played in promoting regionalization is practical: payers wield the greatest leverage in directing surgical care to high-volume providers. A less obvious explanation for the central role of payers, however, is the fact that volume-based referral is an approach that is markedly more attractive to payers than to individual patients and surgeons, and caters much more specifically to their interests.

For payers, volume-based referral is an attractive approach to lowering morbidity and mortality because it is relatively easy to implement. For providers to improve outcomes would require fundamental changes in processes of care, which is a challenging task not only because of the inertia of established practice patterns, but also because of the difficulty identifying which processes need to change. On the other hand, payers can obtain better aggregate outcomes by simply stipulating that specific high-risk surgical procedures are reimbursed only when provided by high-volume providers.

Volume-Based Referral and Patients

The interests of payers are more specifically addressed by volume-based referral than the interests of patients. Although the upside of regionalization (better outcomes and lower costs) benefits both patients and payers, the downside of regionalization (displacement from familiar patterns of healthcare) is borne chiefly by patients. In the patient’s balance, the value of local, familiar care may even outweigh the value of lower operative risk.

Another reason that volume-based referral is more attractive to payers than to patients relates to the fact that the benefit of such a policy is experienced in the aggregate for payers, but individually for patients. Higher volume, being an indirect measure of quality, is not a perfect predictor of better surgical

<table>
<thead>
<tr>
<th>Procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pancreatic resection</td>
</tr>
<tr>
<td>Esophageal resection</td>
</tr>
<tr>
<td>Colorectal resection</td>
</tr>
<tr>
<td>Gastric resection</td>
</tr>
<tr>
<td>Carotid endarterectomy</td>
</tr>
<tr>
<td>Abdominal aortic aneurysm repair</td>
</tr>
<tr>
<td>Lower extremity bypass</td>
</tr>
<tr>
<td>Cerebral aneurysm surgery</td>
</tr>
<tr>
<td>Nephrectomy</td>
</tr>
<tr>
<td>Cystectomy</td>
</tr>
</tbody>
</table>

Table 1. Procedures for Which Multiple Studies Show Evidence for an Association Between Volume and Outcome.2,4
outcomes. If choosing a provider is a gamble, choosing a high-volume provider is simply a gamble with better odds. Nevertheless, it is inevitable that some higher volume providers will deliver consistently low-quality care, and some lower volume providers will deliver high-quality care. For payers, this is an acceptable gamble because better outcomes are measured in the aggregate. For individual patients, however, choosing a provider is a single high stakes gamble. As such, it is much more important to individual patients that high quality providers be identified with greater precision than can be accomplished by simply basing the decision on procedure volume.

For some very high-risk procedures for which the volume–outcome association is strongest, such as pancreatectomy and esophagectomy, the gamble involved in choosing a provider is less of an issue for individual patients (ie, volume is a better predictor). However, for procedures with weaker (but real) volume–outcome associations, the gamble may become more dangerous for patients while remaining a winning strategy for payers. To illustrate this point, one can imagine a scenario in which high-volume providers perform a procedure with an average mortality rate of 4.5% and low-volume providers perform the same procedure with an average mortality of 5.5%. This would roughly correlate with results for a procedure such as elective colectomy in the Medicare population. If one then imagines that (for reasons unrelated to volume) the mortality rates in the high- and low-volume strata are normally distributed around the mean for the group, one would expect to find some overlap in mortality rates, as shown in Figure 1. In the scenario illustrated, about one-quarter of high-volume providers might be expected to perform this procedure with mortality risks that are as high or higher than the mean of the low-volume provider group. Similarly, within the low-volume provider group there would be about one-quarter of providers who perform this procedure with mortality risks at or lower than the mean for the high-volume group. One can easily see how volume-based referral under this scenario might remain a winning strategy for payers, but present a less-than-optimal gamble for individual patients.

**Volume–based Referral and Providers**

Although volume-based referral may be a boon for high volume providers, for those who are lower volume providers (the majority) this payer strategy may appear unfairly discriminatory. Although patient safety is clearly the highest priority, from a purely economic perspective, few providers benefit from regionalization, while many face losses. It is not surprising that many surgeons bristle at the idea of volume-based referral. Although it may be difficult to feel sympathy for those low volume providers who deliver low quality care, there are also likely to be low-volume, high-quality providers who are unfairly bypassed.

Chief among the objections to volume-based referral voiced by low volume surgeons is “I am the exception.” One of the most visible responses to selective referral initiatives such as regionalization is an effort on the part of individual surgeons to track their own results and thereby directly prove the quality of the care they provide. In this effort there has been encouragement from such influential organizations as the Society of American Gastrointestinal and Endoscopic Surgeons and the American College of Surgeons. Although it is true that actual clinical outcomes are intuitively more convincing than a proxy for quality such as volume, the effort to prove quality with outcomes is hopeless for a surgeon in a low-volume practice because of lack of statistical power. In fact, careful examination reveals that even higher volume providers typically do not perform enough procedures to demonstrate statistically that their outcomes are significantly better or worse than national benchmarks. In sum, because of a lack of statistical power, the assumption that a lower volume provider offers care with worse outcomes is an implicit indictment for which even the best surgeons have no easy defense.

**Limitations of Volume–Outcome as a Quality Improvement Measure**

Although for many procedures, higher volume is associated with better outcomes and regionalization could result in lives saved, the practical limitations of this approach must be clearly understood. Specifically, a policy of volume-based referral does not address surgical quality directly, is applicable to only a very small segment of surgical care, and is logistically problematic. Perhaps the most fundamentally concerning aspect of volume-based referral is the way this approach sidesteps the critical issue of surgical quality. Volume-based referral masquerades as a potent surgical quality initiative despite the fact that it does absolutely nothing to improve the quality of care that surgeons deliver, but instead simply moves patients
to providers who are statistically more likely to deliver safer care. Ideally, the central focus of policies surrounding surgical quality should be to improve the way surgeons actually deliver care. Improving surgical quality is a challenging undertaking in which both providers and patients ought to participate, as the interests of both parties are served by it. Cooperative efforts along these lines have begun to emerge, but they are still relatively few and far between.

Volume-based referral as a surgical quality measure is also very limited in the proportion of surgical patients it can affect. The potential impact of a volume-based referral policy is determined by two factors: the strength of the volume–outcome association and how commonly the procedure is performed. As it turns out, there are few high-risk procedures that are both common and have a strong volume–outcome association.

Procedures that are less common and have a strong volume–outcome association (e.g., pancreatectomy, hepatectomy, and esophagectomy) are often considered good candidates for volume-based referral. Because they are less frequently performed, fewer patients would require displacement and, at least in theory, the logistics of regionalization would seem simpler. However, there are two important problems with regionalizing less common high-risk procedures. First, because they are less common, regionalizing them realizes relatively few “lives saved.” Second, their relative scarcity would result in difficulty finding or establishing high-volume regional centers. For example, in the case of pancreatectomy, over 80% of graphically-defined tertiary care markets in the United States do not have a hospital that meets Leapfrog criteria for a high-volume center, and over 60% do not even have enough cases to create a high volume center de novo by concentrating all cases in one hospital.

Procedures with weaker volume–outcome associations that are more common (e.g., colectomy) might be considered good candidates for volume-based referral because of the large number of people who might benefit, even though the benefit is not as large. However, regionalizing these procedures would result in many patients being displaced for smaller individual gains. Because of this, regionalizing these procedures is much less attractive from the perspective of patients. Also, insofar as common procedures are the “bread and butter” of surgical practice, their removal is potentially more disruptive to surgeons’ interests.

Volume-based referral for procedures that are high risk, common, and have significant volume–outcome associations (e.g., cardiac surgery) would potentially result in the greatest number of lives saved. However, these procedures result in high enough numbers of adverse outcomes to allow valid statistical comparisons across providers. In the presence of statistically valid direct comparisons of quality (i.e., outcomes), proxies for quality such as volume have been shown to have less utility.

On closer scrutiny, it would appear that few procedures are ideal candidates for a generalized policy of regionalization. Some of the logistical obstacles might be avoided by more graphically specific, limited applications of volume-based referral. For example, regionalization could be applied only to healthcare markets with existing high-volume providers. Essentially, this is an accommodation that the Leapfrog Group offered by excluding rural areas from their proposed volume-based restrictions. However, the other nonlogistical problems with volume-based referral (e.g., fairness) would still remain unsolved.

Should Volume-Based Referral Be Abandoned?

Volume-based referral falls short because it caters disproportionately to payers’ interests, lacks specificity in identifying better providers, does not directly improve quality of surgical care, and is logistically problematic. Given these many shortcomings, one must ask whether payers should continue to pursue volume-based referral policies. Plainly, payers’ interests are not in volume per se, but in directing patients to better quality, less costly surgical care. Payers are using volume simply because no one has produced a viable alternative metric for judging surgical quality.

Direct measurement of outcomes gets the most attention as a potential alternative to volume as a gauge of quality. Proponents of the National Surgical Quality Improvement Program (NSQIP) have argued that volume should be abandoned as a quality indicator in favor of risk-adjusted, clinical outcomes data. However, the NSQIP remains generally accessible only to larger hospitals. Furthermore, despite a robust data analytic model, it lacks the statistical power to identify quality for specific procedures, and cannot confidently rate the quality of any but the few outlier hospitals. Whether a program such as the NSQIP could be used to predict procedure-specific mortality and morbidity any more accurately than volume-based assessment is doubtful. What is clear is that volume-based referral, despite its many shortcomings, requires no more than counting cases and is therefore significantly easier to apply.

Process-based assessment is another potential alternative to volume as a quality measure. Process-based assessment and improvement is an area with growing momentum, as evident in the popularity of the Surgical Care Improvement Project and the relatively recent use of process measures in pay-for-performance initiatives. Process-based quality measures are appealing because processes of care represent what providers actually do, independent of chance outcomes and patient mix. However, the movement to harness process as a quality lever is still in its infancy. Major challenges to process-based quality assessment include: an insufficient number of discrete processes of care for which there is clear evidence of superiority; the large number of processes that would have to be evaluated to characterize a provider’s quality related to even just one category of procedure; and the formidable logistical challenge of actually measuring compliance with these many processes. Unfortunately, a process-based approach is not ready for prime time.

As long as there remains no viable alternative, one cannot expect payers to abandon volume-based referral policies that
are clearly in their interest. The challenge for surgeon leaders, health services researchers, and payers is to develop methods to measure (and hopefully improve) surgical quality that address the interests of all stakeholders in surgical care. Specifically, the approach must reflect the obligation to ensure patients receive optimal care, fairly identify and reward providers who deliver high quality care, and assure payers that their dollars are well spent. Only with the progress of efforts to establish better, measurable quality standards in surgery will imperfect proxies for quality such as volume begin to fade away. In the meantime, procedure volume will likely continue to be a significant part of the national dialogue surrounding surgical quality.

References